

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:00:14 ; Search time 8809 Seconds  
(without alignments)  
11257.675 Million cell updates/sec

Title: US-10-009-791-21  
 Perfect score: 2288  
 Sequence: 1 gcacgaggccacacattaca.....acattctcaattttgtgaat 2288

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:★

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hug.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_on.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_to.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hg\_hum.\*
- 31: em\_hg\_inv.\*
- 32: em\_hg\_other.\*
- 33: em\_hg\_mus.\*
- 34: em\_hg\_pln.\*
- 35: em\_hg\_rod.\*
- 36: em\_hg\_mam.\*
- 37: em\_hg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hugo\_hum.\*
- 40: em\_hugo\_mus.\*
- 41: em\_hugo\_other.\*

pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C	1	2298	100.0	2288	6	AX046795 Sequence
	2	805.8	35.2	2291	8	AF036340 Arabidops
	3	805.8	35.2	2335	8	AX045625 Arabidops
	4	804.6	35.0	1779	8	AX133556 Arabidops
	5	685.8	30.0	2088	8	AX168845 Oryza sat
	6	683.2	29.9	711	8	AX536527 Glycine m
	7	683	29.9	2163	8	AK121543 Oryza sat
	8	668.4	29.2	2240	6	AX046793 Sequence
	9	666.8	29.1	2305	8	AK101514 Oryza sat
	10	581.6	25.4	2300	8	AK066392 Oryza sat
C	11	581.6	25.4	2324	8	AK100694 Oryza sat
	12	428	18.7	108847	8	AF002109 Arabidops
	13	412	18.0	482	6	AX046779 Sequence
	14	376.8	16.5	134982	8	AF003279 Oryza sat
	15	372.4	16.3	1728	8	AX046791 Sequence
	16	362.6	15.8	126532	8	AC130602 Oryza sat
	17	362.6	15.8	138467	2	AC130812 Oryza sat
	18	300.8	13.1	115270	8	AC135205 Oryza sat
	19	296.6	13.0	844	6	AX046777 Sequence
	20	233.2	10.2	1074	6	AX046801 Sequence
C	21	233.2	10.2	1074	6	BT009618 Triticum
	22	223.4	9.8	794	6	AX046781 Sequence
	23	131.6	5.8	1392	6	AX560749 Sequence
	24	129	5.6	701	6	AX046775 Sequence
	25	119.4	5.2	2646	8	AF139835 Populus t
	26	114.8	5.0	562	6	AF139835 Sequence
	27	113.4	4.9	2147	8	AK100862 Oryza sat
	28	111.4	4.9	2337	8	AF327430 Arabidops
	29	110.6	4.8	134976	8	AP004968 Lotus cor
	30	110.4	4.8	1816	8	BT001946 Arabidops
C	31	110.4	4.8	2225	8	AF005048 Arabidops
	32	107.8	4.7	486	6	AX046799 Sequence
	33	105.8	4.6	2956	6	AK072338 Oryza sat
	34	105.2	4.6	266	6	AR246507 Sequence
	35	97.2	4.2	139855	8	OSJN000554 Arabidops
	36	96.6	4.2	1872	8	BT006340 Arabidops
	37	96.6	4.2	2554	8	AY091772 Arabidops
	38	96.2	4.2	1789	8	AY150427 Arabidops
	39	96.2	4.2	2070	8	AF291816 Arabidops
	40	94.6	4.1	2052	8	AF045799 Arabidops
C	41	94.4	4.1	140234	8	AC016779 Genomic s
	42	94.4	4.1	164899	8	AC084218 Oryza sat
	43	92.8	4.1	2123	8	AK100506 Oryza sat
	44	92.8	4.1	2609	8	AK111900 Oryza sat
	45	92.2	4.0	1801	8	BT002118 Arabidops

## ALIGNMENTS

[illegible]

TITLE		Disease resistance factors	
JOURNAL		Patent: WO 0068406-A 21 16-NOV-2000;	
FEATURES		E.I. DU PONT DE NEMOURS AND COMPANY (US)	
source	Location/Qualifiers		
	1. .2288		
	/organism="Glycine max" /mol_type="unassigned DNA" /db_xref="taxon:3847"		
ORIGIN			
Query Match		100.0%; Score 2288; DB 6; Length 2288;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 2288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GCACGAGGCGACACGTTTACACAGGCGACATATGGTTGCCGGAACAAATCCGGATGGGAA	60
Db	1	GCACGAGGCGACACGTTTACACAGGCGACATATGGTTGCCGGAACAAATCCGGATGGGAA	60
Qy	61	GGGTGTATGTAGTGTCTTCTAGATGAATATTTGTGTATTAACAGAAACGCGGTTTGAAGCAGT	120
Db	61	GGGTGTATGTAGTGTCTTCTAGATGAATATTTGTGTATTAACAGAAACGCGGTTTGAAGCAGT	120
Qy	121	GACGTGTATCATCAGTACATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	180
Db	121	GACGTGTATCATCAGTACATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	180
Qy	181	AGTTTGT	240
Db	181	AGTTTGT	240
Qy	241	AACCCATTCCTTTTGGCACTTGAACCAAAACCTCTGCACCTTTTCTTTTCACTCTCAG	300
Db	241	AACCCATTCCTTTTGGCACTTGAACCAAAACCTCTGCACCTTTTCTTTTCACTCTCAG	300
Qy	301	TCTCCGATCCATATGACGAGGAAACGACGTCGCGGACACACGTCGTGTGTGTGTGTGTGTGT	360
Db	301	TCTCCGATCCATATGACGAGGAAACGACGTCGCGGACACACGTCGTGTGTGTGTGTGTGTGT	360
Qy	361	CTCGACTGCTCATCCCTTACATCGACACCCCAAGGACCGGACCGGTTTCCCAGGT	420
Db	361	CTCGACTGCTCATCCCTTACATCGACACCCCAAGGACCGGACCGGTTTCCCAGGT	420
Qy	421	GTGTGACGCTGT	480
Db	421	GTGTGACGCTGT	480
Qy	481	CTACACCAACCCCGGCTCGCTCGCGCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCT	540
Db	481	CTACACCAACCCCGGCTCGCTCGCGCGCGCTTCCGCGCTTCCGCGCTTCCGCGCTTCCGCGCT	540
Qy	541	CAAGGCGAAGCCCGGAGCGCAATGTTTCACTTTGATACCCGAGGATTTGGGCGGACACGT	600
Db	541	CAAGGCGAAGCCCGGAGCGCAATGTTTCACTTTGATACCCGAGGATTTGGGCGGACACGT	600
Qy	601	CATCTCCGTGGGTCAAGAGATTTCTCAGTACTTGTGATGCTCAAGAGCCTTCCACTTCCG	660
Db	601	CATCTCCGTGGGTCAAGAGATTTCTCAGTACTTGTGATGCTCAAGAGCCTTCCACTTCCG	660
Qy	661	CCGATGATTTCAAGGATTTCCGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTT	720
Db	661	CCGATGATTTCAAGGATTTCCGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTT	720
Qy	721	TCAGGCTCTCAAGCTTGACAGTGTCTCCGTTTCCACCGATGCTTCTTTTCCATATCCG	780
Db	721	TCAGGCTCTCAAGCTTGACAGTGTCTCCGTTTCCACCGATGCTTCTTTTCCATATCCG	780
Qy	781	TCGCTTTTTCAGAGTTTAAAGTCTTGTTTTGGAGGAAGCTCAATTTCTTGAGAGGA	840
Db	781	TCGCTTTTTCAGAGTTTAAAGTCTTGTTTTGGAGGAAGCTCAATTTCTTGAGAGGA	840
Qy	841	CGGGAATGGCTACACGAGCTTGTGTGATTAATACAGTCTTCTGAGACTCTCAATTTTTTA	900
Db	841	CGGGAATGGCTACACGAGCTTGTGTGATTAATACAGTCTTCTGAGACTCTCAATTTTTTA	900
Qy	901	CTTGACAGACATTTGCTGTGTGAAGATTTGAGACCTTGAACTTTTAGCTTAAATAATTTGCC	960
Db	901	CTTGACAGACATTTGCTGTGTGAAGATTTGAGACCTTGAACTTTTAGCTTAAATAATTTGCC	960
Qy	961	CAACTTAGTGTGTGAACACTTACTGACTGTGAATACTGGATCTTGTGAACCTTTCTTTAA	1020
Db	961	CAACTTAGTGTGTGAACACTTACTGACTGTGAATACTGGATCTTGTGAACCTTTCTTTAA	1020
Qy	1021	GCATGCCCTCTGCGCTGGAAGAGTTTGTGGAGGCACTTACACAGGAGACCAAGAGATA	1080
Db	1021	GCATGCCCTCTGCGCTGGAAGAGTTTGTGGAGGCACTTACACAGGAGACCAAGAGATA	1080
Qy	1081	CTCTCTATATCATTTACAGCAAAAGTTATGTCTGTAATCTGGATCTTGTGAAGAA	1140
Db	1081	CTCTCTATATCATTTACAGCAAAAGTTATGTCTGTAATCTGGATCTTGTGAAGAA	1140
Qy	1141	TGAGTTGCCCATTTGTGTCTATGTTTGCAGCCGCTACTTAAATAATTTGGATCTCTATGC	1200
Db	1141	TGAGTTGCCCATTTGTGTCTATGTTTGCAGCCGCTACTTAAATAATTTGGATCTCTATGC	1200
Qy	1201	AATGTGTACACGAGGATCATTTGTATTAATCCAAAGGTGTCAAATCTCGAAAGTCT	1260
Db	1201	AATGTGTACACGAGGATCATTTGTATGTTAATCCAAAGGTGTCAAATCTCGAAAGTCT	1260
Qy	1261	TGAGCAAGGATTAATTTGGAGATAGAGGTTAGAGGTTCTTGTGCTGTTGTGTGAAG	1320
Db	1261	TGAGCAAGGATTAATTTGGAGATAGAGGTTAGAGGTTCTTGTGCTGTTGTGTGAAG	1320
Qy	1321	GCTTAAAGGCTTAGGATTTGAAAGGCGCATCATCATCAAGGAATGGAGATGAAGAAG	1380
Db	1321	GCTTAAAGGCTTAGGATTTGAAAGGCGCATCATCATCAAGGAATGGAGATGAAGAAG	1380
Qy	1381	TACTGTGTCCCATAGAGGCTAATAGCCCTGTGTCAGAGGCTTTCAGAGCTTGAATACAT	1440
Db	1381	TACTGTGTCCCATAGAGGCTAATAGCCCTGTGTCAGAGGCTTTCAGAGCTTGAATACAT	1440
Qy	1441	GGCTGTTTATGTCGTGATTAATTAATGATCTCTGGAACATATTTGGAACCTCAGTTGA	1500
Db	1441	GGCTGTTTATGTCGTGATTAATTAATGATCTCTGGAACATATTTGGAACCTCAGTTGA	1500
Qy	1501	GAACTCTGTGATTTTCGCTTGTGTCATCAAGGAAGAGATTAACATGATTTGCT	1560
Db	1501	GAACTCTGTGATTTTCGCTTGTGTCATCAAGGAAGAGATTAACATGATTTGCT	1560
Qy	1561	ACTTGCAATGGGTGAGGCTCTACTAGGGGCTGTGCAAGCTGAGGAGATTTGCTCT	1620
Db	1561	ACTTGCAATGGGTGAGGCTCTACTAGGGGCTGTGCAAGCTGAGGAGATTTGCTCT	1620
Qy	1621	ATATCTCAGGCTGGGGGCTGATGATGAGGCTTGTGTCATTTGGAACATACAGTCC	1680
Db	1621	ATATCTCAGGCTGGGGGCTGATGATGAGGCTTGTGTCATTTGGAACATACAGTCC	1680
Qy	1681	AAATGTGAGATGGATCTCTCTGTTTATGTTGGGGAGTCTGATGAGGCTTTTGGAGTT	1740
Db	1681	AAATGTGAGATGGATCTCTCTGTTTATGTTGGGGAGTCTGATGAGGCTTTTGGAGTT	1740
Qy	1741	CGTAAAGGGGTGTCCTAGTCTTCAGAAACCTTGAATGAGAGGCTTTTATTTTTCAGTGA	1800
Db	1741	CGTAAAGGGGTGTCCTAGTCTTCAGAAACCTTGAATGAGAGGCTTTTATTTTTCAGTGA	1800
Qy	1801	ACGTGACATGCTGCTGGCTGCAACAAATTTGATCTTCTTAGTCTTGTGGGTGCAAGG	1860
Db	1801	ACGTGACATGCTGCTGGCTGCAACAAATTTGATCTTCTTAGTCTTGTGGGTGCAAGG	1860
Qy	1861	TATGCTGTATCTCCATCTGGACGTGATCTTTTGGTAAATGGCTCGACCTTTTGGAAACAT	1920
Db	1861	TATGCTGTATCTCCATCTGGACGTGATCTTTTGGTAAATGGCTCGACCTTTTGGAAACAT	1920
Qy	1921	TGAGTTGATTTCTTTAGAAAGGTGCTACGAATACAATCCAGATGAGACTGTAGTTGT	1980
Db	1921	TGAGTTGATTTCTTTAGAAAGGTGCTACGAATACAATCCAGATGAGACTGTAGTTGT	1980

```

1981 TGAGATCCTGCTCATATTTCTTGATATTTATTTCTTGCAGGCGAGATCAGATTTTCC 2040
Db 1981 TGAGATCCTGCTCATATTTCTTGATATTTATTTCTTGCAGGCGAGATCAGATTTTCC 2040
Qy 2041 AGATACTTGTGCTTGGACACTGCCACATGCTTGATACCTAGAGCCAGAGCTGTG 2100
Db 2041 AGATACTTGTGCTTGGACACTGCCACATGCTTGATACCTAGAGCCAGAGCTGTG 2100
Qy 2101 TATATATACAGTGTCTTTGTTTCTTCTTCCCTTTCATATGCTGTTTCTATGTTCC 2160
Db 2101 TATATATACAGTGTCTTTGTTTCTTCTTCCCTTTCATATGCTGTTTCTATGTTCC 2160
Qy 2161 TGCTCTATTGTAGTTCATTTTGTAGCAATTTAGTCTTGTAAAGCTGTTTCTATTG 2220
Db 2161 TGCTCTATTGTAGTTCATTTTGTAGCAATTTAGTCTTGTAAAGCTGTTTCTATTG 2220
Qy 2221 AAATTCGAAAGCGTCCCTTAACGCTATTGGCTCCCTTAAAACTGAACATTTCAATT 2280
Db 2221 AAATTCGAAAGCGTCCCTTAACGCTATTGGCTCCCTTAAAACTGAACATTTCAATT 2280
Qy 2281 TTGTGAAT 2288
Db 2281 TTGTGAAT 2288

RESULT 2
AF036340 2291 bp mRNA linear PLN 28-MAY-1998
LOCUS Arabidopsis thaliana LRR-containing F-box protein (COI1) mRNA,
DEFINITION complete cds.
ACCESSION AF036340
VERSION AF036340.1 GI:3158393
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Feyé,B.J., Benedetti,C.S., Penfold,C.N. and Turner,J.G.
AUTHORS Arabidopsis mutants selected for resistance to the phytotoxin
TITLE coronatine are male sterile, insensitive to methyl jasmonate, and
resistant to a bacterial pathogen
JOURNAL Plant Cell 6, 751-759 (1994)
REFERENCE 2 (bases 1 to 2291)
AUTHORS Xie,D.X., Feys,B.F., James,S., Nieto-Rostro,M. and Turner,J.G.
TITLE COI1: an Arabidopsis gene required for jasmonate-regulated defense
and fertility
JOURNAL Science 280 (5366), 1091-1094 (1998)
MEDLINE 98248619
PUBMED 9582125
REFERENCE 3 (bases 1 to 2291)
AUTHORS Turner,J.G., Xie,D., James,S., Feys,B.J. and Nieto-Rostro,M.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1997) School of Biological Sciences, University
of East Anglia, University Plain, Norwich, Norfolk NR4 7TU, England
FEATURES
source
1. .2291
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/map="between as and cer-8"
1. .2291
/gene="COI1"
/note="genomic sequence submitted as GenBank Accession
Number AF002109; Original gene name T29M12.10"
263. .2041
/gene="COI1"
/codon_start=1
/product="LRR-containing F-box protein"
/protein_id="AAC17498.1"

```

```

/db_xref="GI:3158394"
/translation="MEDEIKRCKLSVAVDDVIEQVMTYITDPKDRDSASLVCRVR
FKIDSEVHVWTKALCVTATPDRSLRPPNLSLKAGKPRAMFNLIPEWNGGVYTP
WVTEISNRLQKLSVHFRTMTVDLDRLAKRADDELTKLCKSGFTDGLISIV
THCKIKITLMEESFSEKQKWLHQAHTNSLEVLNFMVTEFAKISPKDLETARN
CRSLVSKVGDFFELVGFKAANLEEFCCGSLNEDIEMPEKYNLFRKLCRLG
LSYMPNEMILIFPFAAQIRKLDLYALLEDEHCTLIQKPNLELVNFRVIGRGL
EVLAAQCKQLKRLIERGADQMGDEGLVQRGLIALAQCGQLEFMAVYVSDITN
ESLSEICTYLNKLCDFRLVLLDREERTDPLDNGVRSLLIGCKLRFAPYLRQGL
TDLGLSIVGYQSPNVRWMLLYGVGSEDEGLMEFSGRCNLOKLEWRGCGCFERAAAA
VTKPUSIRYLMVQGRASMTQDMQMARPNWIELIPSRVPEVNVQOGEIREMEHPA
HILAYSLAQGRDTCPTTVRVLKEPI"

```

## ORIGIN

```

Query Match 35.2%; Score 805.8; DB 8; Length 2291;
Best Local Similarity 66.9%; Pred. No. 1.4e-179;
Matches 1196; Conservative 0; Mismatches 577; Indels 14; Gaps 3;
Qy 349 GGTGACGCTGCTCGATCGCTCATCCCTTACATCGACGACCCCAAGAGCCGCGACGC 408
Db 310 GGTGATGATGTCATCGAGCAAGTCATGACCTATATAACTGACCCGAAAGATCGCGATT 369
Qy 409 CGTTTCCAGGTGTGTCGACGCTGGTACAGCTGACTCGCTCACCCGCAACGACGTCCAC 468
Db 370 GGCCTCTTTGGTGTGTCGAGATGGTTCAAGATTGATTCCGAGACGAGAGCATGTGTAC 429
Qy 469 CATCGCGCTCTGTACACACACCCGCGCTCGCTCCGCGCGCTTCCCGCACTCCGA 528
Db 430 TATGCGCTTTGTCTACACTGACGCCCTGATGCTGTAGCGCTGATTTCCGAACTTGAG 489
Qy 529 GTCGCTCAAGCTCAAGGCGAAGCCCGGACCGCAATGTTCAACTTGAATACCCGAGATTG 588
Db 490 GTGCTCAAGCTTAAAGCGAAGCCTAGACGACTATGTTTATCTGATCCCTGAGAACTG 549
Qy 589 GGGCGGACGCTCACTCCCTGGGTCAAGAGATTCTCAGTACTTCGATTGCTCAAGAC 648
Db 550 GGGAGGTTATGTTACTCTCTGGGTTACTGAGATTCTTAAACCTTAGCAGCTCAATC 609
Qy 649 CCTCCACTTCCGCGCATGATTGTCAAGGATTCGATCTTCAAGATCTCGCTCGTGACCG 708
Db 610 GGTGACCTTCCAGCGAGTATGTGACGTACTAGATCTAGATCGTTTAGCTAAAGCTAG 669
Qy 709 CGGTCAAGCTGCTTCAAGCTCTCAAGTTGACAAGTGTCCGGTTTCCACCGATGGTCT 768
Db 670 AGCAGATGATCTTGAGACTTTGAGCTAGACAAAGTGTCTGCTTTTACTACTGATGGACT 729
Qy 769 TTTCCATATCGTCTGCTTTTGCAAGTTTAAAGTCTTGTGTTTGGAGGAAAGCTCAAT 828
Db 730 TTTGAGCATCGTTACACACTGCGAGGAAATAAAACCTTTGTTAATGGAGAGAGTCTTT 789
Qy 829 TCTTGAGAGGACGAGAGATGGCTACACAGCTTGTCTTTGAATAATACAGTTCTTGGAC 888
Db 790 TAGTGAAAGGATGTTAGTGGCTTTCATGAGCTTGTCTAGCAGACACACATCTCTTGA 849
Qy 889 TCTCAATTTTACTTGACAGACATGCTGCTTGTGAAGATGAGGACCTTGAACCTTTAGC 948
Db 850 TTTAAACTTCTACATGACGAGTGTGCAAAATCAGTCCCAAAAGATCTTGGAAACCAATAG 909
Qy 949 TAAATATGCCCCCAACTTAGTCTGTGAAACTTACTGACTGTGAATACTGATCTTGT 1008
Db 910 TAGAAATGCGCGCTCTCTGGTATCTGTGAAGGTGCGTGTGATTTGAGATTTTGAACATAG 969
Qy 1009 GAACCTTCTTTAAGCATGCTCTGCGCTGGAAGAGTGTGTTGGAGGACCTTCAACAGGAG 1068
Db 970 TGGGTTCTTTAAGGCTGACGCTAACTTGAAGATTTTGTGGTGGCTCTTGTGAATGAGGA 1029
Qy 1069 -----ACCAGAAAGATATCTGCTATATCATTTACAGCAAAAGATTATGTCGATGG 1119
Db 1030 TATTGGAATGCTGAGAGTACATGAATCTGGTGTGTTTCCCGGAAACTATGTCGGCTGG 1089
Qy 1120 TTTAATATATATGGAAGATGAGTTGCCCATTTGTTTCATGTTTGGAGCGCTTACTAAA 1179
Db 1090 TCTCTCTTATGAGGACCTAATGAATGCCATATCTATTTCCATTCGCGGCCCAATCCG 1149

```

Qy	1180	AAAAATTGGATCTCTCTATGCAATGCTAGACACGAGGATCATTTGTAATGTTAAATCCAAAG	1233
Db	1150	AAAGCTGGAATTTGCTTTATGTCATTGCTAGAAACTGAAGACCCATTGTAAGCTTATCCAAAA	1209
Qy	1240	GTGTCCAAATCTGGAAGTCTTTGAGACAAAGGAATGTAATTTGGAGATAGAGGTTAGAGGT	1299
Db	1210	GTGTCCCTAAATTTGGAAGTCTTCGAGACAAGGAATGTAATCGGAGATAGGGGCTTAGAGGT	1269
Qy	1300	TCTTTGGTCTGTTTGTAAAGGCTTAAAAAGGCTTAGGATTTGAAAGGGCGCATGATGATCA	1359
Db	1270	CCTTGCACAGTACTGTAAAGCAATTTGAAGCGGCTGAGGATTTGAAGCGGTCGAGATGAACA	1329
Qy	1360	AGGAATGGAGATGAAGAAGGATCTGTGTCCCATAGAGGGCTAATAGCCTTTGTACAGGG	1419
Db	1330	AGGAATGGAGACGAAGAAGGCTTTAGTCTCACAAAGAGAGATTAAATCGCTTTGGCTCAGGG	1389
Qy	1420	CTGTTCAGAGCTTCAATACATGCGCTGTTTATGCTGTGATATTAACAATGCATCTCTGGA	1479
Db	1390	CTGCCAGAGCTAGAAATACATGGCGGTGTATGTCCTCAGATATACTAACGAATCTCTTTGA	1449
Qy	1480	ACATATTGGAACCTCACTTTGAAGAACCTCTCTGTGATTTTCGCGCTTGTTGCTTGACCATGA	1539
Db	1450	AAGCATAGGCACATATCTGAAAAACCTCTGTGACTTCGCGCTTGCTTTACTCGACCGGGA	1509
Qy	1540	AGAGAAATAACTCAATTTGCCACTTGACAAATGGGGTAGGGCTCTACTAGAGGGGCTGTGA	1599
Db	1510	AGAAAGATTTACAGATCTGCCACTTGGACACGGAAGTCCGATCTCTTTTGATTTGGATGCAA	1569
Qy	1600	CAAGCTCAGGAGATTGCTCTATATCTCAGCGGTGGCGGTTGACTCATGTAGGCCCTTGG	1659
Db	1570	GAAACTCAGACGATTTCGATCTTATCTTGAGCAAGCGCGCTTAACCGACTTGGGCTTAAG	1629
Qy	1660	TTACATTGGACAATACAGTCCAAATGTGAGATGGATCTGCTTGTTTATGTGGGGGAGTC	1719
Db	1630	CTACATCCGACAGTACAGTCCAAACGTTGAGATGGATGCTGCTGGGTTACGTAGGTGAATC	1689
Qy	1720	TGATCAGGCGCTTTTGGAGTTCGTAAGGGGTGCTTAGTCTTCAGAAAACTTGAATGAG	1779
Db	1690	AGATGAAGGTTTAATTTGAAATTTCTCAAGAGGCGTCCAAATCTACAGAAAGCTAGAGATGAG	1749
Qy	1780	AGGCTGTATTATTTTTCAGTGAACGTGCACATGCTGGCTGGCTGCAACAAATGACTTCTCT	1839
Db	1750	AGTTTGT---TGCCTTCACTGAGCAGACAAATGCGCTGAGCGGTTACAAAAATGGCTTCACCT	1806
Qy	1840	TAGGTACTTTGTGGGTGCAAGGTTATGTTGATCTCCCATCTGGACGTGATCTTTTGGTAAAT	1899
Db	1807	GAGATACTTTGTGGGTACAAGGTTACAGAGCATCCATGACCGGGAACAAGATCTAATTCGAGAT	1866
Qy	1900	GGCTCGACCCCTTTTGGAAACATTCGAGTTGATTTCCCTTCTAGAAAGGTGGCTACGAATACCA	1959
Db	1867	GGCTAGACCGTACTGGAACATCGAGCTGATTCATCAAGAAAGAGTCCCGGAAGTGAATCA	1926
Qy	1960	TCCAGATGAGACTGTAGTTGTTGAGCATCTCTGCTCATATTTCTTGCAATATTATTTCTTTGC	2019
Db	1927	ACAAAGGAGATGAAGAGATGGAGCATCCGGCTCATATATGGCTTACTACTCTCTGCGC	1986
Qy	2020	AGGCAGAGATCAGATTTTCAGATACCTGTTTGGCTTTTGGACACTGCCACATGGCGTTGA	2079
Db	1987	TGGCCAGAGAACAGATTTGTCCAAACAATGTTTAGAGTCTCTGAAGGAGCCAATATGATATG-	2045
Qy	2080	TACCTAGAGCCGACGCTGTGATATATACCAAGTTTCTTTTCTTTT 2126	
Db	2046	-ACCCAAAAACAGGTTTGTATATAAAGATTTTGTAGTCTCGAGTTTT 2091	
RESULT 3			
AY045625			
LOCUS	AY045625	2335 bp	mRNA
DEFINITION	Arabidopsis thaliana	At2g39940/t28M21.10	mRNA, complete cde.
ACCESSION	AY045625		
VERSION	AY045625.1		
KEYWORDS	FLI CDNA.		



3'UTR		HILAYSLAGQRTDCTPTVRVLKEP1"		1980..2335	
ORIGIN		Query Match		35.2%; Score 805.8; DB 8; Length 2335;	
		Best Local Similarity		66.9%; Pred. No. 1.4e-179;	
		Matches 1196; Conservative		0; Mismatches 577; Indels 14; Gaps 3;	
QY	349	GGTCGACGTGGTCTCGACATCGCTCATCCCTTACATFCGACGACCCCAAGAGCCGCGACGC	408		
Db	248	GGTTGATGATGATCATCGACGAAGTCATGACCTATATTAACGTACCCGAAAGATCGCGAATTC	307		
QY	409	CGTTTCCAGGTGTGTGACGCTGGTGAAGCTCGACCTCGCTCACCCGACGACGTCCAC	468		
Db	308	GGCTTCTTTGGTGTGTGCGAGATGGTTCAAGATTGATTCGAGACGAGAGAGATGTGAC	367		
QY	469	CATCGCGTCTCTACACACACCCCGCTCGCTCCGCGCGCTTCCCGCACTCGA	528		
Db	368	TATGCGCTTTGCTACATGCGACGCTGATCGTCTTAGCGCTCGATTCGCCGAACCTTGAG	427		
QY	529	GTGCTCAAGCTCAAGGCAAGCCCGACCGCCGCAATGTTCAACTTGATACCCGAGATGG	588		
Db	428	GTGCTCAAGCTTAAAGGCAAGCTTAGACGAGCTATGTTAATCTGATCCCTGAGAACTG	487		
QY	589	GGGCGACACGTCACTCCCTGGGTCAAGAGATTTCTAGTACTTCTGATTCGCTCAAGAG	648		
Db	488	GGGAGGTTATGTTACTCTCTGGGTACTAGATTTCTTAAACACCTTTAGCGAGCTCAATC	547		
QY	649	CTTCACTTCCCGCATGATGTCAAGATTCCGATCTTCAAGATCTCGCTCGTGACCG	708		
Db	548	GGTGCACTTCCGACGATGATGTGATGACCTAGATCTAGATCGTTTAGCTAAAGCTAG	607		
QY	709	CGGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG	768		
Db	608	AGCAGATGATCTTGAACCTTGAAGCTAGACAAAGTCTTCTGCTTTTACTACTGATGACT	667		
QY	769	TTTCCATATCGTGGCTTTTGAAGATTTAGAGTCTTGTGTTTGGAGGAAGCTCAAT	828		
Db	668	TTTGACATCGTTTACACATCGCAGGAAATATAAACTTTGTTAATGGAAGAGATTTCTT	727		
QY	829	TCTTGAGAGGAGGAGATGCTCACAGCTTGTGTTTGAATAATACAGTCTTCTGACAG	888		
Db	728	TAGTGAAGAGGATGGAAGTGGCTTCATGAGCTTGTCTGACACACACATCTCTTGAGT	787		
QY	889	TCTCAATTTTACTTGACACATCTGCTGTGTGAAGATTTGAGGACCTTGAACCTTTTAC	948		
Db	788	TTTAAACTTCTACATCAGCGAGTTTGCACAAATCAGTCCCAAGACTTGGAAACCATAGC	847		
QY	949	TAAATTTGCCCACTTAGTCTGTGAACTTACTGACTGTGAAATCTGATCTTGT	1008		
Db	848	TAGAAATTCGCGTCTCTGTGATCTGTGAAGGTTCGGTGAATTTGAGATTTTGAACATG	907		
QY	1009	GAACTTCTTTAAGCATGCTCTCGCTTGAAGAGTTTGTGAGGAGCACCTTACACAGGAG	1068		
Db	908	TGGTTTCTTTAAGGCTTGACCTAATCTTGAGAAATTTTGTGGTGGCTTCTTGATGAGA	967		
QY	1069	-----ACGAGAAAGATCTCTGCTATATCATTAACAGCAAAAGTTATGTGATTTGG	1119		
Db	968	TATTTGGAATTCCTTGAGAGTACATGAATCTGTTTTCCTCCGAAACTATGTGCGCTTG	1027		
QY	1120	TTTAAACATATTTGGAAGATGATGTTGCCATTTGTTTCATCTTTTCCAGCCGCTACTAA	1179		
Db	1028	TCTCTCTTACATGGACCTAATGAAATGCCAATCTATTTTCCATTCGCGGCCCAATCCG	1087		
QY	1180	AAATTTGATCTCTCTATGCAATGCTPAGACGAGGATCATTTGATGTTTAAATCCAAAG	1239		
Db	1088	AAAGTGGATTTGCTTTATGCAATGCTTGAACCTGAAGACATTTAGCGTTTATCCAAA	1147		
QY	1240	GTGTCCAAATCTGGAAGTCTCTGAGACAAAGGAATGTTAATTTGAGATAGAGGTTTAGG	1299		
Db	1148	GTGTCTTAAATTTGGAAGTCTCTGAGACAAAGGAATGTTAATTTGAGATAGAGGTTTAG	1207		
QY	1300	TCTTGGTGTGTTTGAAGGCTTAAAGGCTTAGGATTTGAAGGGGCGATGATGATCA	1359		
Db	1208	CCTTGCACTACTGTAAAGCAGTTGAAGCGCTGAGGATTTGAAGCGGTGCGATGAACA	1267		
QY	1360	AGGAATGAGGATGAAGAGGTTACTGTGTCCCATAGAGGCTTAATAGCTTTGTACAGGG	1419		
Db	1268	AGGAATGAGGACGAGAGGCTTAGTCTCTCAAAAGAGATTTAAATCGCTTTGGCTCAGGG	1327		
QY	1420	CTGTTTCAGAGCTTGAATCATGCTGTTTATGCTGCTGATATTAACAAATGCAATCTCTGA	1479		
Db	1328	CTGCAGGAGCTAGAAATACATGCGCGGTGTATGCTCAGATATTAACAAATCTCTTGA	1387		
QY	1480	ACATATTCGAACTCACTTGAAGAACCTCTGTGATTTTCCCTTTGTGTTGCTTGACCATGA	1539		
Db	1388	AGCATAGGACATATCTGAAAGAACCTCTGTGACTTCCGCTTTGCTTACTCGACGGGA	1447		
QY	1540	AGAGAAGATACTGATTTGCCACTTGACAAATGGGGTGAAGGCTCTACTGAGGGCTTGA	1599		
Db	1448	AGAAAGGATTAACAGATCTGCCACTCGACAAACGAGAGTCCGATCTCTTTGATTTGGATG	1507		
QY	1600	CAAGCTGAGGAGATTTGCTCTATATCTCAGCGCTGCGGGTTGACTGATGTAGGCTTGG	1659		
Db	1508	GAACTCAGACGATTTGCACTTATCTGAGACAAAGCGGCTTAAACGACTTGGGCTTAAG	1567		
QY	1660	TTACATTCGACAAATCAGTCCAAATGTGAGATGGAATGCTTGTGTTATGTGGGGAGTC	1719		
Db	1568	CTACATCGACAGTACAGTCCAAACGTGAGATGGAATGCTGCTGGTTAGTAGTGAATC	1627		
QY	1720	TGATCAGGGCTTTTGGAGTTTCGCTTAAGGGTGTCTTAGTCTTTCAGAACTTGAATGAG	1779		
Db	1628	AGATGAAGTTTAAATGGAATTTCTCAAGAGGCTGTCCAAATCTACAGAGCTTAGAGTAG	1687		
QY	1780	AGGGTGTGTTTATTTTTCAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1839		
Db	1688	AGGTTGT---TGCCTTTCAGTGAAGCAATCTGCGAGGGTTTCAAAATTTGCCCTTCACT	1744		
QY	1840	TAGGTACTTGTGGGTGCAAGGTTATGTTGTTATCTCCATCTGGACGCTGATCTTTTGGTAAT	1899		
Db	1745	GAGATCTTGTGGGTACAGGTTTACAGATCTGATGAGCGGACCAAGATCTAATGCAGAT	1804		
QY	1900	GGCTCGACCTTTTGGAACTTGAATGATTCCTTCTAGAAAGGTGGCTACGATACCA	1959		
Db	1805	GGCTAGACGCTTGGAACTGAGCTGATTCATCAAGAGAGTCCCGAAGTGAATCA	1864		
QY	1960	TCCAGATGAGCTGATGTTGTTGAGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2019		
Db	1865	ACAAAGAGAGATTAAGAGATGGAGATCCGGCTCATATATTGGCTTACTTCTCTGGC	1924		
QY	2020	AGGGCAGAGATCAGATTTTCCAGATCTGTTGCTTTCGACACTTGGACACTGCCCATGCGTTGA	2079		
Db	1925	TGGCCAGAGAACAGATTTGTCACCACTGTTAGAGTCTCTGAGGAGCCCAATATGATATG-	1983		
QY	2080	TACCTAGAGCCAGAGCTGTGATATATATACAGTCTTTCTTTTGTGTTT 2126			
Db	1984	-ACCCAAAACAGGTTTGTATATAAGATTTTGTAGTCTCGAGTTT 2029			
RESULT 4					
AY133556					
LOCUS					
DEFINITION					
Arabidopsis thaliana At2g39940/T28M21.10 mRNA, complete cds.					
ACCESSION					
AY133556					
VERSION					
AY133556.1					
KEYWORDS					
FLI CDNA.					
SOURCE					
Arabidopsis thaliana (thale cress)					
ORGANISM					
Arabidopsis thaliana					
REFERENCE					
AUTHORS					
Chen, R., Chen, H., Kim, C.J., Shinn, P., Ban, J., Bowser, J.,					
Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,					
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,					
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,					



QY 1480 ACATATTGGAACCTCCTGTAATTTTCGCTTGTGTTGCTTGACCATGA 1539  
 Db 1488 AAGCATAGGCACATATCTGAAAAACCTCTGTGACTTCGCGCTTGTCTTACTCGACCGGGA 1247  
 QY 1540 AGAAGAATAACTGATTTGCCACTTGACAAATGGGGTGAAGGCTCTACTGAGGGGCTGTGA 1599  
 Db 1248 AGAAGGATACAGATCTGCCACTGGAACAACGAGTCCGATCTCTTTGATTGGATGCA 1307  
 QY 1600 CAAGCTGAGGAGATTGCTCTATATCTCAGGCGTGGCGGGTGTGACTGATGTAGGCTTGG 1659  
 Db 1308 GAAACTCAGACGATTGCTATCTATCTGAGACAAGCGCGCTTAACCGACTTGGGCTTAAG 1367  
 QY 1660 TTACATTGGCAATACAGTCCAAATGTGAGATGGATGCTGCTTGTGTATGTGGGGAGTC 1719  
 Db 1368 CTACATCGACACGTACAGTCCAACTGAGATGGATGCTGCTGGTACTAGTAGGTGATC 1427  
 QY 1720 TGATGACGGCTTTTGGAGTTGCTTAAGGGTGTCTAGTCTTTCAGAACTTGAATGAG 1779  
 Db 1428 AGATGAAGGTTTAAATGGAATCTCAAGAGGCTGTCCAAATCTACAGAACTAGAGATGAG 1487  
 QY 1780 AGGGTGTATTTTTCAGTGAAGCTGCACTTGTGCTGTGCTGCAACACAAATGACTTCTCT 1839  
 Db 1488 AGGTTGT---TGCTTCAGTGAAGCAATCTGCTGAGCGGTTACAAAATTGCTTCTACT 1544  
 QY 1840 TAGGTACTTGTGGTGCAAGGTTATGTTGTTATCTCATCTGACGCTGATCTTTTGTGTAAT 1899  
 Db 1545 GAGATACTTGTGGGTACAGGTTACAGATCGATGAGCGGACAGATCTAATGCAGAT 1604  
 QY 1900 GGCTGACCCCTTTTGGAACTTGAAGTTGATTCCTTTAGAAAAGTGGCTACGAATACCA 1959  
 Db 1605 GCTAGACCGTACTGGAACATCGAGCTGATTCATCAAGAGAGTCCCGGAAGTGAATCA 1664  
 QY 1960 TCCAGATGAGCTGATGTTGTTGAGATCTCTGCTCATATCTTGCATATATCTTCTTCTG 2019  
 Db 1665 ACAAGGAGAGATAAGAGATGAGCATCCGGCTCATATATTGGCTTACTACTCTCTGTC 1724  
 QY 2020 AGGCGAGATCAGATTTTCCAGATCTGTT 2050  
 Db 1725 TGGCCAGAGAAACAGATTGTCCAACTGTT 1755

RESULT 5  
 AY168645  
 LOCUS Oryza sativa (japonica cultivar-group) COI1 mRNA, complete cds.  
 DEFINITION  
 ACCESSION AY168645  
 VERSION  
 KEYWORDS  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 1 (bases 1 to 2088)  
 Wang, W. and Wang, X.  
 Direct Submission  
 TITLE Submitted (25-OCT-2002) Biochemistry, Fudan University, 220 Handan  
 JOURNAL Road, Shanghai 200433, P.R. China  
 FEATURES  
 source  
 1. .2088  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="japonica"  
 /db\_xref="taxon:39947"  
 /chromosome="1"  
 66. .1853  
 /note="rCOI1; similar to Arabidopsis thaliana COI1"  
 /codon\_start=1  
 /product="COI1"  
 /protein\_id="AA038719.1"  
 /db\_xref="GI:37359393"  
 /translations="MGVEPPEPRRLNRALSFDWVPEALHLVGMHVEDPPREAAAR  
 VCRWRHIDALTRKHVTVAFYAAARPARLRFRPRLSLKPKPRAAYGLIPDDWG

CDS

ORIGIN  
 Query Match 30.0%; Score 685.8; DB 8; Length 2088;  
 Best Local Similarity 63.7%; Pred. No. 3.6e-151;  
 Matches 1111; Conservative 0; Mismatches 617; Indels 15; Gaps 4;  
 QY 345 GTGTGGTGCAGTGTGCTCGACTGCGTCACTCCCTTACATGACAGACCCCAAGGACCGCG 404  
 Db 124 GGGTCCCCGAGCGCGCTGCACCTCTGTGATGGGCCAGTGCAGGACCCGGGACAGG 183  
 QY 405 AGCGCGTTTCCAGAGTGTGCGAGCTGGTACGAGCTCGACTCGCTCAACCCCAAGCAGC 464  
 Db 184 AGCGCGCTCGCGGGTGTGCGCGCTGCGACCGCATCGACGCTCACGCGCAAGCAGC 243  
 QY 465 TCACATCGCGCTCTGTCTACACCAACCCCGGCTCGCTCGCGCGCGCTTCCCGCAC 524  
 Db 244 TCACGCTCGCTTCTGTCTACGCGCGCGCGCGCTTCCGGAGCGGTTCCCGCGGC 303  
 QY 525 TCGAGTCCCTCAAGCTCAAGGCGAGCGCGCAATGTTCAACTTGATACCGGAG 584  
 Db 304 TCGAGTCTGCTCTGCTCAAGGCGAGCGCGCGCGCATGATCGAGCTCGCGCGCTCGAGTGCCTCA 363  
 QY 585 ATTGGGGCGGACACCTCACTCCCTGGGTCAAGAGAGATTTCTCAGTACTTTCGATTGCTCA 644  
 Db 364 ACTGGGGCGCTACCGCGCGCATGATCGAGCTCGCGCGCGCTCGAGTGCCTCA 423  
 QY 645 AGAGCTTCCACTTCCGCGCATGATGTCAAGATTCGATCTTCAGATCTCGCTCGTG 704  
 Db 424 AGGCGCTCCACCTCGCGCGCATGACCGTCAACGCGCGCATCGCGCGCTTGTCCGCG 483  
 QY 705 ACCGGGTGACGTGCTTCAAGCTCTCAAGCTTGAAGTGTCCGTTTCAACCGAGT 764  
 Db 484 CCGCGGACACATGCTGCGAGGAGCTCAAGCTCGAAGTGCATCGGCTTCTCCAGTGC 543  
 QY 765 GTCTTTTCCATATCGCTGCTTTTCAAGAGTTTAAGAGTCTTGTTTTGGAGGAAAGT 824  
 Db 544 CCTCGGCTCTCGCGCGCTCGTGCAGATCCCTGAGAACTTTATTTCTGGAAGAGTGC 603  
 QY 825 CAATTCTTGAAGGACGAGAAATGGCTACAGAGCTTGTCTTGAATAATACAGTCTTG 884  
 Db 604 ATATTACTGATAAGGGTGGTGAATGGCTTCATGAACCTTGTCTGTCACAAATCTGTTCTG 663  
 QY 885 AGACTCTCAATTTTACTTGCAGACATGCTGTTGTGAAGATTGAGGACCTTGAACTTT 944  
 Db 664 TGACACTGAACTTCTACATGACTGAACCT---CAAGTGGCGGCGAGCTGATCAGAGCTTC 720  
 QY 945 TAGCTAAAATAATGCCCCCAACTTAGTGTCTGTGAACCTTACTGACTGTGAAAATCTGGATC 1004  
 Db 721 TTGCAAGAGATTGCAAGTCAATGATTTCAATGAAGATGATGAGTGTGATCTTTCAGATC 780  
 QY 1005 TTGTGAACCTTTTAAAGCATGCTCTCGCTCGAAGAGTTTGTGAGGACCACTTCAACG 1064  
 Db 781 TGATTAGTTTTTTTCAACAGCAATGCGCTGCAAGACTTGTCTGAGAGAGCAATCTCAG 840  
 QY 1065 AGGAACCCGA-----AAGATCTCTGTATATCATTTACAGCAAGTATGTGAGTTGG 1118  
 Db 841 AGGTAGGAGAGCTCACAAAGTATGAAAAGTTAAGTTCCACCCAGATATGCTTCTTGG 900  
 QY 1119 GTTTAAACATATATGAAAAGATGAGTTGCCCAATGTGTTCATGTTTGAGCGGCTACTAA 1178  
 Db 901 GGTCTTACCTACATGGGAACAATGAGATGCTGCTTATCTTCCCTTTTTCGATGAAACTCA 960  
 QY 1179 AAAAAATGATCTCTCTATGCAATGCTAGACACGAGATCATTTGATGTATTAATCAA 1238

961	AGAAACTGGAC	TTGCAATACACTTTTCTCACAACAGAAAGATCATTTGTCAGATTATTGCAA	1020
1239	GGTGCCAAATCTGGAAGTCCCTTGAGACAAGGAATGTAATTCGAGATAGACGGTTAGAGG	1298	
1021	AATGTCCTCAATCTACCTAATCTTTGAGTGAGGAACGTGATAGAGATAGAGGGCTAGAAG	1080	
1299	TTCTTGGTTCGTTGTTGTAAGAGGCTAAAAAGCGTTAGGAATGAAAGGGCGATGATGATC	1358	
1081	TTGTTGGTGATACATGCAAGAAGCTACGAAGACTCCGAATTCGACGGGCTGATGATGATC	1140	
1359	AAGGAATGGAGGATGAAGAAGGTTACTGTGTCCCATAGAGGGCTAATAGCCTTTGTCACAGG	1418	
1141	CAGGTCGACGGAAGACGACGAGGAGGATTTCTCAGCTAGGCTTGACAGCGCTGCTGTTG	1200	
1419	GCTGTTTCAGAGCTTGGAATACATCGCTGTTTATCTGTCTGATATTACAAATGCATCTCTGG	1478	
1201	GTTCGCGTGAAATGGAGTACATAGCTGCCATGCTATCGGATATCACCAATGGGCGCTTGG	1260	
1479	AAACATATTGGAACCTCACTTCGAAGAACCTCTGTGATTTTCGCCCTTGTTGCTTTGACCATG	1538	
1261	AGTCTATTGGGACTTTCTGCAAAAATCTATACGACTTTTGGCTTTGTGCTACTTGAACAGAG	1320	
1539	AAGAGAAGATAACTGATTGGCCACTTGCAAAATGGGGTGAGGCTCTACTGAGGGGCTGTG	1598	
1321	AAAGACAGGTAAACAGATCTGCCACTTGACAAATGGTGTCTGTGCTGTTTAAGAAAATTGCA	1380	
1599	ACAGCTGAGGAGATTGCTCTATCTCAGGGGTGGCGGTTGACTGATGTAAGGCTTGG	1658	
1381	CNAAGCTTCGAGGTTTGCTCTCTACCTTTAGACACGAGGGCTTTGAGATGAGGCTTTA	1440	
1659	GTATACATTGGACAATACAGTCCMAATGTGAGATGGATGCTTGCTTGGTTATGTGGGGAGT	1718	
1441	GCTACATCGGACAGTACAGTGGAAATATCCAATACATGCTACTGGGCAATGTTGGTGAAT	1500	
1719	CTGATGACGGGCTTTTGAGGTTGCTGAAGGGGTGCTCTAGTCTTTCAGAAAATTTGAAATGA	1778	
1501	CTGACCATGGATTGATCCGTTTTCGACGTGGGCTGCACCAACCTTTCAGAAGCTTGAATTTGA	1560	
1779	GAGGGTGTTATTTTTCAGTGAAACGTGCACCTTGCTGTGGCTGCAACAACATTCGACTTCTC	1838	
1561	GAAGCTG---CTGCTTCAGCGAGCGAGCTTTGTCCCTCGCTGTACTCGCAGTGCCTTCCC	1617	
1839	TTAGTACTTCTGGGTCAAGGTTATGGTGTGATCTCCATCTGGACGCTGATCTTTTGGTAA	1898	
1618	TGAGATACATATGGGTGCAAGATACAGAGCATCTCAAACAGGCGCTTGACCTCTCGTCA	1677	
1899	TGGCTCGACCCCTTTTGGAAACATGAGTTGATTCCTCTTAGAAAGGTGGCTACGAATACCA	1958	
1678	TGGCCAGGCCCTTTCTGGAAACATCGAGTTTACACCTCCGAGCCCTGAGAGTTTAAATCAT	1737	
1959	ATCCAGATGAGACTGTGA---GTTGTTGAGCATCTCTGCTCATATTCTTTGCAATTATTCTC	2015	
1738	TGACAGAAAGATGGAGAACCCCTGTGTGGATAGCCATGCTCAGGTTCTTGGCTACTATTCCC	1797	
2016	TTGAGGGCAGAGATCAGATTTTCCAGATACTGTTGTGCTTTTGGACACATGCCATGCG	2075	
1798	TTGCTGGGAAGAGGCTCTGACTGCCCTCAGTGGGTGATCCCTTGTGATCCTCCGTTGTT	1857	
2076	TTG	2078	
1858	TTG	1860	

RESULT 6	AF536527	711 bp	DNA	linear	PLN 01-OCT-2002
LOCUS	AF536527				
DEFINITION	Glycine max putative coronatine-insensitive 1 (CoRI) gene, partial cds.				
ACCESSION	AF536527				
VERSION	AF536527.1	GI:23395152			
KEYWORDS					
SOURCE	Glycine max (soybean)				
ORGANISM	Glycine max				

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

1 (bases 1 to 711)  
Christiansen,L.C. and Ulvskov,P.  
Putative Glycine max sbCOI1 homologous to Arabidopsis COI (atFBL2)  
Unpublished

2 (bases 1 to 711)  
Christiansen,L.C. and Ulvskov,P.  
Direct Submission  
Submitted (08-AUG-2002) Biotechnology Group, Danish Institute of  
Agricultural Sciences, Thorvaldsensvej 40, Frederiksberg C 1871,  
Denmark

Location/Qualifiers

1. .711  
/organism="Glycine max"  
/mol\_type="genomic DNA"  
/cultivar="TGGX1835-2E"  
/db\_xref="taxon:3847"  
<1. .>711  
/gene="COI1"  
/note="sbCOI1"  
<1. .>711  
/gene="COI1"  
/product="putative coronatine-insensitive 1"  
<1. .>711  
/gene="COI1"  
/note="sbCOI1; similar to Arabidopsis COI1 (atFBL2) "  
/codon\_start=1  
/product="putative coronatine-insensitive 1"  
/protein\_id="AA3N31713.1"  
/db\_xref="GI:23395153"  
/translation="SSMQAVSRHGLIALSQGCELYMAVYVDITNALSLEHIGTHLK  
NLCDRLVLLDHEEKITDPLDNGVRALRGCDKLRFPALYLRGGLTVDVGLYIGQY  
SNPVRWMLGYVGE DAGLEFAKGCPSLQKLEMRGCLFFSERALAAVATQLTSLRYL  
WVGQYGVSPSGDLLVYMWAPFWMNIELIPSRKVATNTNPDET VVVEHPAHLIAYISLWG  
ORSDRPTVVEVLDATMLD"

FEATURES

source

gene

mRNA

CDS

## ORIGIN

Query Match	29.9%;	Score 583.2;	DB 8;	Length 711;
Best Local Similarity	99.6%;	Pred. No. 1.2e-150;		
Matches 685;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1383	CTGTGTCACATAGAGGGCTAAATAGCCCTTGTCAGAGGCTGTTCAGAGCTTGAATCATGG	1442	
DB	14	CGGTGCCCATAGAGGGCTAAATAGCCTTGTCAGAGGCTGTTCAGAGCTTGAATCATGG	73	
QY	1443	CTGTTTATGTGCTGATATTACAATTCATCTCTGGAACATATTGGAACCTCATTGGAAGA	1502	
DB	74	CTGTTTATGTGCTGATATTACAATTCATCTCTGGAACATATTGGAACCTCATTGGAAGA	133	
QY	1503	ACCTCTGTGATTTTGGCCTTGCTGTGCTGACCATGAAGAAGATAACTGATTTGCCAC	1562	
DB	134	ACCTCTGTGATTTTGGCCTTGCTGTGCTGACCATGAAGAAGATAACTGATTTGCCAC	193	
QY	1563	TTGCAATGGGGTGAGGGCTCTACTAGGGGCTGTGCAAGCTTGAGAGATTTGCTCTAT	1622	
DB	194	TTGCAATGGGGTGAGGGCTCTACTAGGGGCTGTGCAAGCTTGAGAGATTTGCTCTAT	253	
QY	1623	ATCTCAGGCTGGGGGTGACCTGATGTAGGCCCTTGCTTACATTTGACATAACAGTCCAA	1682	
DB	254	ATCTCAGGCTGGGGGTGACCTGATGTAGGCCCTTGCTTACATTTGACATAACAGTCCAA	313	
QY	1683	ATGTGAGATGGATGCTGCTTTGGTTATGTGGGGAGTCTGATGCAAGGCTTTTGGAGTTCG	1742	
DB	314	ATGTGAGATGGATGCTGCTTTGGTTATGTGGGGAGTCTGATGCAAGGCTTTTGGAGTTCG	373	
QY	1743	CTAAGGAGTGTCTTAGTCTTCAGAACTTCGAAATGAGAGGGTGTATTATTTTTCAGTGAAC	1802	
DB	374	CTAAGGAGTGTCTTAGTCTTCAGAACTTCGAAATGAGAGGGTGTATTATTTTTCAGTGAAC	433	
QY	1803	GTGCACCTTGCTGTGGCTGCACACAAATTTGACTTCTCTTATGACTCTTGCGGTGCAAGTT	1862	





QY	585	ATTGGGCGGACACGTCACTCCCTGGGTCAAAGAGATTCTCAGTACTTCGATTGGCTCA	644
DB	380	ACTGGGCGCTACGCGCGCATGGATCGACGAGCTCGCGCGCGCTCGATGCTCTCA	439
QY	645	AGAGCTCCACTCCGCGCGCATGATTCTCAAGGATTCGGATCTTCAGAATCTCGCTCGTG	704
DB	440	AGGCGCTCCACCTCCGCGCGCATGACCGTACCGAGCGCGCATCGCGCCCTTGTCGCGG	499
QY	705	ACCGCGGTACGTCTTCACGCTCTCAAGCTGTGAAGTGTCCGGTTTCCACACCGATG	764
DB	500	CCCGGGACACATGCTGCAGGAGCTCAAGCTCGACAAGTGCATCGGCTCTCCACTGACG	559
QY	765	GTCCTTTCCATATCGGTGCTTTTGCAGAGTCTTTAAGAGTCTTGTTTTTGGAGGAAAGCT	824
DB	560	CCCTCCGCTCGTCCGCGCTCGTGCGAGTCCCTGAGAACTTTATTTCTGGAAGAGTGCC	619
QY	825	CAATCTCTGAGAAGCAGGAGAATGGCTACACGAGCTTGCTTTGAAATTAATACAGTCTTG	884
DB	620	ATATTACTGATGAAGGTTGGTGAATGCTTCATGAACCTTGCTCAACAATTTCTGTTCTGG	679
QY	885	AGACTCTCAATTTTATCTGACAGACATGCTGTTGTGAAGATTTGAGGACCTTGAACCTT	944
DB	680	TGACACTGAATCTTACATGACTGGACT---CAAGTGGCGCGAGCTGATCTAGAGCTTC	736
QY	945	TAGCTAAATTTGCCCACTTAGTGTCTGTGAACCTTACTGACTGTGAAATCTATGATCT	1004
DB	737	TTGCAAGAATTTGCAAGTCATTTGATTTTCAATGAGATGAGTGTGATCTTTTCAGATC	796
QY	1005	TTGTGAATCTCTTTAAGCATGCTCTCGCTGGAAGAGTTTGTGAGGACCTACCAAG	1064
DB	797	TGATTAGTTTTTTTCAAACAGCAATGCGCTGCAAGACTTTGCTGGAGGAGCAATCTACG	856
QY	1065	AGGAACCCAGA-----AAGATCTCTGTATATCATTTACCAGCAAAAGTTATGTCTGATGG	1118
DB	857	AGTGTAGGAGAGCTCACCAAGTATGAAAGTTAAGTTTCCACCGCGATATGCTTTCTGG	916
QY	1119	GTTTACATATATTGGAAGAATGAGTTGCCATTTGTTTCATGTTTTCAGCCGCTACTAA	1178
DB	917	GGCTTACTCTACATGGGAACAAATGAGATGCTGTATCTTCCCTTTTCGATGAAACTCA	976
QY	1179	AAAAATTTGGATCTCCTCTATGCAATGCTAGCACGGAGGATCATTTGTTAATCCAA	1238
DB	977	AGAACTCTGGACTTGCATATACATTTTCTCAACAAGAGATCATTTGCAGATATTGCAA	1036
QY	1239	GGTGTCCAAATCTCGAAGTCTCTGAGACAGGAATGTAATTTGGAGATAGAGGGTTAGAGG	1298
DB	1037	AATGTCCAAATCTACTAATTTCTGAGGTGAGAACGTGATGAGGATAGAGGGCTAGAAG	1096
QY	1299	TTCTTGCTCGTTTGTGAAGAGCTTAAAGAGCTTAGGATTAAGAGGGCGGATGATGATC	1358
DB	1097	TGTTGTGTGATACATGCAAGAGCTACGAGACTCCGAATTTGAGCGGGTGTATGATGATC	1156
QY	1359	AAGGAATGGAGGATGAAGAAGTATGTGTCCATAGAGGGCTATAGCTTGTACACAGG	1418
DB	1157	CAGGCTGCGAGGAGACAGGAGGAGTTTCTCAGCTAGGCTTGAACGCCGTGCTGTG	1216
QY	1419	GCTGTTTCAGAGCTTGAATACATGCGCTGTTTATGTGCTGATATTAACAATGCAATCTCTGG	1478
DB	1217	GTTTCCGCTGAATTTGGAGTACATAGCTGCGCTATGATCGGATATCAACAATGGGCGCTGG	1276
QY	1479	ABACATTTGAACTCACTTGAAGAACTCTGTGAATTTTCGCTTGTGTTGCTTGACCATG	1538
DB	1277	AGTCTATTGGGACTTTCTGCAAAAATCTATACGACTTTGGCTTGTGCTACTGAGGGCTGTG	1598
QY	1539	AAGAGAGATPAACTGATTTGGCATCTTGACATCGGGGTGAGGCTCTACTGAGGGCTGTG	1598
DB	1337	AAAGACAGGTAACAGATCTGCCACTTGACAATGGTGTCTGTGCTCTGTTAAGAAATGCA	1396
QY	1599	ACAAGCTGAGGAGATTTGCTCTATATCTCAGGGGTGCGGGTTCAGTGTATGAGGCTTG	1658
DB	1397	CAAGCTTCGAGGTTTGCTCTCTACCTTAGACAGGAGGCTTTCAGATGATGAGGCTTA	1456

QY	1659	CTTACATGGGCAATACAGTCCAAATGTGCAGATGGAGTCTGTTGGTTATTGTGGGGAGT	1718
Db	1457	GCTACATCGGACGTACAGTGGAATATCCAATACATGCTACTGGGCAATGTTGGTGAAT	1516
QY	1719	CTGATCAGCGGCTTTTGGAGTTTCGCTAAGGGGTGCCCTAGTCTTCAGAAAACTTGAATGA	1778
Db	1517	CTCACCATGGATTCGCTTTTCGAGTGGGCTGCACCACCTTCAGAAGCTTTGAATTGA	1576
QY	1779	GAGGGTGTATTTTTTCAGTGAACGTGCACCTTCTGTGTGGCTGCACAATAATTGACTTCTC	1838
Db	1577	GAAGCTG---CTGCTTCAGGACGGAGCTTTTGTCCCTGCTACTGTGCAGATGCGCTCCC	1633
QY	1839	TTAGGTACTTTTGGGTGCAAGGTTATGGTGTATCTCCATCTGACGTGATCTTTTGGTAA	1898
Db	1634	TGATATACATATGGTTCAGAGTACAGAGCATCTCAAACAGCCCTTGACCTCCCTCTCA	1693
QY	1899	TGGCTCAGCCCTTTTGGAACTTTGAGTTGATTCCTTCTAGAAAAGTGGGTACGAATACCA	1958
Db	1694	TGGCCAGCGCTTCTTGAAACATCGAGTTTACACTCCGAGCCCTGAGAGTTTAAATCAT	1753
QY	1959	ATCCAGATGAGACTGTGA---GTTTGTGAGCATCTCTGCTCATATCTTGCATATATTCTTC	2015
Db	1754	TGACAGAGATGGAGAACCTTGTTGTGATGGCCATGCTCAGGTTCTTGGCTACTATATCCC	1813
QY	2016	TTGCAGGGCAGAGATCAGATTTTCCAGATCTGTTGTGCCCTTTGGACACTGCCACAT	2072
Db	1814	TTGCTGGAAGGAGGTCTGACTGCCCTCAGTGGGTGATCCCCCTTGCACTCTCGCTGAT	1870
RESULT 8			
LOCUS AX046793 2240 bp DNA linear PAT 15-DEC-2000			
DEFINITION Sequence 19 from Patent WO0068406.			
ACCESSION AX046793			
VERSION AX046793.1 GI:11876305			
KEYWORDS Oryza sativa			
SOURCE Oryza sativa			
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Ephemerophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Berrharctoideae; Oryzaceae; Oryza.			
REFERENCE 1 Gaimi, P.G., Pamodu, O.O., Lee, J.M., Miao, G.H. and Maxwell, C.A.			
AUTHORS Disease resistance factors			
TITLE Patient: WO 0068406-A 19 16-NOV-2000;			
JOURNAL E.I. DU PONT DE NEMOURS AND COMPANY (US)			
FEATURES			
source Location/Qualifiers			
1..2240 /organism="Oryza sativa"			
/mol_type="unassigned DNA"			
/db_xref="taxon:4530"			
ORIGIN			
Best Match 29.2%; Score 668.4; DB 6; Length 2240;			
Query Local Similarity 63.4%; Pred. No. 4.7e-147; Indels 15; Gaps 4			
Matches 1093; Conservative 0; Mismatches 616;			
QY	353	GACGTGTCTCGACTGGCTCATCCCTTACATCGACGACCCCAAGACCGCGHACGCCGTT	412
Db	290	GAGGAGCGCTGCACTTGGCTGTGGGTCTGGGTACGTGACGACCCCGGACGAGGGAGCGGTG	349
QY	413	TCCCAGGTGTGTGACAGCTGTGTACGAGCTGCACTCGGTCTACCCGACGACGTCACCATC	472
Db	350	TGCTCTGTGTCGCGCGCTGGCACCGCATCGACGCGCTCAGCGGAGACGTCACCGTG	409
QY	473	GCGTCTGCTACCAACACCCCGGCTCGCTCCGCGCGCTTCCCGACCTCGAGTCG	532
Db	410	CCCTTCTGTCTAGCGCGCTGCGCGCGGACCTGTCTCGCGGTTCCCGCTGGAGTCG	469
QY	533	CTCAGCTCMAGGGCAAGCCCGCAGCGCCAAATGTTCAATTGATACCGAGGATTTGGGC	599
Db	470	CTCGCGGTTCAGGGAGCGCGCGCCCATGTACGGGCTCATCCCGAGGACTGGGCG	528



593 GGACAGTCACTCCCTGGGTCAGAGAGATTTCTCAGTACTTTCGATTGCTTCAAGAGCCTC 652  
 530 GCTACGCGCGCCCTGGTGGCGGAGCTCGCGCGCCGCTCAGTGCCTCAAGGCGCTC 589  
 653 CACTTCGCGCGCATGATTTGTCAGAGATTCGATCTTCAGAAATCTCGCTCGTGAACCGGCT 712  
 590 CACTCGCGCGCATGATTTGTCAGAGATTCGATCTTCAGAAATCTCGCTCGTGAACCGGCT 649  
 713 CAGTGTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 772  
 650 CACATGCTGAGAGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 709  
 773 CATATCGCTCGCTTTCAGAGATTTAAGAGTCTTGTGTTTGGAGGAAAGCTCAATTTCTT 832  
 710 CTGCTGCGCGCTCTCAGATCACTGAGAACATTTATTTCTGGAGGAATGCTCAATTTCT 769  
 833 GAGAGGACGAGAAATGGCTACAGAGCTTGTGTTTGAATTAATACAGTCTTGAAGCTCTC 892  
 770 GATATGCTACTGAATGGCTCCAGAGCTTGTGTTTGAATTAATACAGTCTTGAAGCTCTC 829  
 893 AATTTTCTTACAGAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 952  
 830 AACTTCCACATGACGAGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 886  
 953 AATTTGCCCCAATTTAGTGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1012  
 887 AAGTGAAGTCACTAAATTTCTGAGAGTCACTGAGTCACTGAGTCACTGAGTCACTGAG 946  
 1013 TTTCTTTAAGCATGCTCTGCGCTGGAAGAGATTTTGTGAGAGCCTTCAAGAGGAACTCA 1072  
 947 TTTTTCGGATGCTGATCACTTTCAGAGAGTTCGCGGAGGAGGATTCATTTGAGCAAGG 1006  
 1073 GA-----AAGATCTCTGCTATATCACTACAGCAAGTATTCGATGCGTTGTTAACTA 1126  
 1007 GAGTCACTAAGTATGAAATGTTAAATTTCCCTTCAAGACTGTGCTCTTGAAGCTTTAGC 1066  
 1127 TATATTGGAAGATGATGTTCCCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1186  
 1067 TACATGGGCAACAGAGATGCCATTTATTTCTTCTTCTGCTTCTGCTTCTGCTTCTGCT 1126  
 1187 GATCTCTCTAAGCAATGCTAGACAGGAGATCAATTTGTTGTTGTTGTTGTTGTTGTTGTT 1246  
 1127 GACTTGCAGTACACTTTTCTCACCACCTGAAGTCACTGCCAATCATTTGCAAAATGTCCT 1186  
 1247 AATCTGGAAGTCTTTGAGCAAGGATGTTAATGAGATAGAGGTTAGAGGTTCTTGGT 1306  
 1187 AACTTACTAGTCTTTCGCTGAGGAATGTTGTTGAGATAGAGATAGAGATAGAGATAGAG 1246  
 1307 CGTTGTTGTAAGAGCTTAAAGGCTTAGGATTTGAAAGGGGCGATGATGATCAAGGAATG 1366  
 1247 GACACATGCAAGAGCTTCAAGAGCTCAGAGTTCAGAGGAGGAGATGATGATCCAGGTTG 1306  
 1367 GAGATGAGAGAGTACTGTGTTCCATAGAGGCTTAAGCTTCTCAGAGGCTGTTTCA 1426  
 1307 CAAGAAGAACAGGAGGAGTCTCTCAAGTCCGTTTGAACAATGTTAGCTAGGATGCGCT 1366  
 1427 GAGCTTGAATCATGCTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1486  
 1367 GAATGGAATCATAGCTGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1426  
 1487 GGAACTCACTTCAAGAACTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1546  
 1427 GGGACTTCTGCAAAAATCTTTTGGACCTTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1486  
 1547 ATAACTGATTTCCCATCTGACAAATGGGTTGAGGCTCTACTGAGGCTGTGACAGCTG 1606  
 1487 ATAACTGATTTCCCATCTGACAAATGGGTTGAGGCTCTACTGAGGCTGTGACAGCTG 1546  
 1607 AGGAGTTTGTCTATATCTCAGGCTGCGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1666  
 1547 CCGAGGTTTGTCTATATCTCAGGCTGCGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1606  
 1667 GGACAATACAGTCCAAATGTGAGATGGATGCTGCTTGGTTTATGTGGGGGAGTCTGATGCA 1726

1607 GGACAGTACAGTGAATTTATCAATACATGCTTCTGGTAAATGTTGGGAAAACAGATGAT 1666  
 1727 GGGCTTTTGGAGTTGCTAAGGGTGTCTAGTCTTTCAGAAAATTTGAAATGAGAGGTTGT 1786  
 1667 GGTCTGATCGGTTTGCATTTGGGTTGTGAGAACCTTCGGGAAGCTTGA--GCTAAGGAGT 1723  
 1787 TTAATTTTTCAGTGAACGTTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1846  
 1724 TGTGTTCTCAGTGAAGCTTTAGCCCGGCTTATAGCGAGTATGCTTCCCTCGAGATAC 1783  
 1847 TTTGCGGTCAGAGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1906  
 1784 GTGTGGGTACAGGCTTCAAGGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 1843  
 1907 CCTTTTGGAACTTCAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1966  
 1844 CCTTTTGGAACTTCAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1903  
 1967 GAGACTGTA---GTTGTTGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2023  
 1904 GATGTTGAACCTTGTGTAGATAGTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 1963  
 2024 CAGAGATCAATTTCCAGATACATGTTGTTGCTTGGCTTGGACACTGC 2067  
 1964 AAGAGGTCGAGCTGCCACGATCTGTGTTGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 2007

## RESULT 9

AKI01514

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

Koya, S., Kurihara, C., Kuroaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Nishikawa, J., Nishi, K., Nomura, K., Numasaka, R., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Ry, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., and Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.  
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Nishikawa, J., Oka, M., Ry, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaka, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayaahizaki, Y.

Location/Qualifiers  
1. 2305  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033046F05"

FEATURES  
source  
1. 2305  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033046F05"

ORIGIN  
Query Match 29.1%; Score 666.8; DB 8; Length 2305;  
Best Local Similarity 63.3%; Pred. No. 1.1e-146;  
Matches 1092; Conservative 0; Mismatches 617; Indels 15; Gaps 4;  
353 GAGCTGTCCTGAGTGGTTCATCCCTTACATCGACGACCCCAAGGACCGGCGCGT 412  
287 GAGGAGGCGCTGCACTGCTGGGTGACGTGGACGACCCCGGCGGAGGCGCGTG 346  
413 TCCGAGGTGTGTCAGCGTGTGACGCTCGACTCGCTCACCGCAAGCAGCTCACCATC 472  
347 TCGCTGCTGTCGCGCGCTGGCAACGCAATCGACGGCTCACGGGAGCACTCACCGTG 406  
473 GCGCTCTGCTACACCAACCCCGCTCGCTCGCGCGCGCTTCCGCACTCGAGTCG 532  
407 CCTTCTGCTACGCGCGCTGCGCGCGCACCTGCTCGCGCGCTTCCGCGGCTGAGTCG 466  
533 CTCAGCTCAAGGGCAAGCCCGGACCGCAATGTTCAACTGATACCGGAGATTGGGCG 592

467 CTCGCGGTCAAGGGGAGAGCGCGCGCCCAATGTAAGGGCTCATCCCGAGAGCTGGGGC 526  
593 GGACACGTCATCCCTGGGTCAAGAGAGATTTCTCAGTACTTTCGATTTGCTCAAGAGCTTC 652  
527 GCCTACGCGCGCGCTGGGTGCGCGAGCTCGCGCGCGCTCGAGTGTCTCAAGGCGCTC 586  
653 CACTTCGCGCGCATGATTTGTCAGAGATTCGATCTTCAAGATCTCGCTGTCGACCGCGT 712  
587 CACTTCGCGCGCATGATTTGTCAGAGATTCGATCTTCAAGATCTCGCTGTCGACCGCGT 646  
713 CAGCTGCTTCAGCTCTCAAGCTTCAAGAGTCTCGCGTTCACCAACGATGCTCTTTTC 772  
647 CACATGCTGCGAGAGCTCAAGCTTCAGAGTCTCGCGTTCACCAACGATGCTCTTTTC 706  
773 CATATCGCTGCTGCTTTTGAAGAGTTTAAGAGTCTTTTGAAGAGTCAAGTCAATCTTT 832  
707 CTCGTCGCGCGCTCTCGCAGATCACTGAGAACATTTATTTCTGAGAGATGCTCAATTTG 766  
833 GAGAGGAGCGAGAGATGCTACAGAGTCTGCTTTGAATATACAGTCTTCTGAGACTCTC 892  
767 GATATGCTGCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 826  
893 AATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952  
827 AACTTCCATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883  
953 AATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1012  
884 AAGTCAAGTCACTAATTTCAATGAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943  
1013 TTTCTTAAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072  
944 TTTTCCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1003  
1073 GA-----AAGATPACTCTGCTATATCATTTACCAGCAAGTTATGTCGATTTGGTTTAA 1126  
1004 GAGCTCAGTAAAGTATGAGTAAATTTAAATTTCCCTTCAAGAGTGTGCTCTTAGGACTTAC 1063  
1127 TATTTGGAAGAAATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1186  
1064 TACATGGGAGCAAAACGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1123  
1187 GATCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246  
1124 GACTTGGAGTACACTTTCTTCCATCTGAGATCACTGCTGCTGCTGCTGCTGCTGCTGCT 1183  
1247 AATCTGGAAGTCTTGGAGCAAGAAATGTAATTTGAGATAGAGGTTAGAGGTTCTTGGT 1306  
1184 AACTTACTAGTCTTGGGTTGAGGATGATTTGAGATAGAGGATTTAGGGGTTGTTGCA 1243  
1307 CGTTGTTGTAAGAGGCTTAAAGGCTTAGAGTTGAAAGGCGCTGATGATCAAGGAATG 1366  
1244 GACACATGCAAGAGCTTACAAAGACTCAGAGTTGAGCGAGGATGATGATCAAGGTTTG 1303  
1367 GAGGTAAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1426  
1304 CAAGAGAACAGAGAGGCTCTCTAGTGGGTTGACACTGTAGCGTAGGATCGGT 1363  
1427 GAGCTTGAATACATGCTGTTTATGTTGATGATTAATAATGATCTCTGGAACATATT 1486  
1364 GAATGGAATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423  
1487 GGAATCACTTGAAGAGCTCTGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1546  
1424 GGGACTTTCTGCAAAAATCTTTGCGACTTCGCTTGTCTGCTGCTGCTGCTGCTGCTGCTG 1483  
1547 ATAACTGATTTGCACTGAGAGTGGGCTGAGGCTCTTACTGAGGGGCTGTGCAAGCTG 1606  
1484 ATAACTGATTTGCACTGAGAGTGGGCTGAGGCTCTTACTGAGGGGCTGTGCAAGCTG 1543  
1607 AGGAGATTGCTCTATATCTCAGGCGTGGGCTGAGTCTGATGATGAGGCTTGGTTATTT 1666



298 GCATGGCGGTA CTCCACACACGCCGCGCTGTTTCAGCGGTTCCCTGCTCCAGTCG 357  
533 CTCAGAGCTCAAGGGCAAGCCCGAGCGCAATGTTCACTTGTATACCCGAGGATGGGCG 592  
358 CTCAGAGCTCAAGGGCAAGCCCGAGCGCAATGTTCACTTGTATACCCGAGGATGGGCG 417  
593 GACAGCTCACTCCCTGGGTCAAGAGATTTCTAGTACTTGGATGCTCAAGAGCTC 652  
418 GGTCCGCTCGCGGTGATCGCCAGCTCTCGGCTCTGCTCCACTTCTCAAGGCGCTC 477  
653 CACTTCGCGCGATGATCTCAAGGATTCGAGATCTCCGATCTCCGCTCGTCCAGCGCGT 712  
478 CACTTCGCGAGATGATCTGTCGAGAGAGCTCGAGCTCGTCTCGTCCGCGCGAGGCG 537  
713 CAGCTGCTTCAAGCTCTCAAGCTTGAAGTGTCTCGGTTTTCACACCGATGGTCTTTTC 772  
538 CACATGCTCTCCCTGTTCAAGCTTGAAGTGTCTCGGTTTTCACACCGATGGTCTTTTC 597  
773 CATATCGGTGCTTTTGAAGATTTGAAGTGTCTCGGTTTTCACACCGATGGTCTTTTC 832  
598 CTGCTCGCGCGCTGCAAGAACTTGAAGTGTCTCGGTTTTCACACCGATGGTCTTTTC 657  
833 GAGAGG---ACGAGAGTGGCTTACACGAGCTTGTCTTGAATTAATACAGTCTTTCAGACT 889  
658 GAGAAAGAAATGACGAATGGATTCGGAGCTTGTCTACCAACATAGTGTCTTTCAGACA 717  
890 CTCATTTTCTTACTTGACAGATTTGCTGTGTGAAGATTTGAGACCTTGAACCTTTTACT 949  
718 CTGAATTTCTTTTTCAGATCTCAGGCGATCCCCA---GCATATCTTACCTCCCTTGTG 774  
950 AAAAAATGCCCCAATCTAGTGTCTGGAATCTGCTGCTGGAATCTGGAATCTGCTG 1009  
775 CGCAATTTGCGAAGCTGAAAGTGTGAAGATTTGAAGATTTGAGTGTCTTCACTGCTG 834  
1010 AACTTCTTTAAGCATGCTCTGCGCTGGAAGATTTTGTGGAGCACCTTACACAGAGAA 1069  
835 GATTTGTTCCGTACAGAGAAATTTACAGACTTTTGGCGGTGTTTCCCTTGTATGATCAG 894  
1070 CCAGAAAGATCTCTGTATA-----TCATTACAGCAAGTTATGTCGA 1114  
895 GGTCAAGTTGAAGAGAGTAGAAATTTATGAATACTACTATTTCCTCTCTCAATTGCTCCGC 954  
1115 TTGGTTTAAATATTTGAAGATGAGTGTCCCATGTTGTTTCATGTTTTCAGCGGTA 1174  
955 TTGAGTTTGTCTTACATGGGACAAAGAGATGCAAGTGTCTATTTCCATATGTTGCTGCA 1014  
1175 CTAATAAAATTTGATCTCTCTATGCAATCTAGACAGGAGATCAATGTTATGTTAATC 1234  
1015 CTCAGAGTTGATCTTCAATTTCAATTTCCATCTTCCACAGAGATCACTGTGCTGCTC 1074  
1235 CAAAGGTGTCATCTGGAAGTCTTTCAGACAGGATTTGATGAGATGAGGTTA 1294  
1075 CAGCGCTGCCAAATCTGGAATCTTAGAGGTGAGGATGTCATAGGGGATCTGGGTTA 1134  
1295 GAGTTCTTCTGCTGTTTGAAGAGCTTAAAGGCTTAGGATTTGAAGGCGGATGAT 1354  
1135 GAGTTCTTCTGCTGTTTGAAGAGCTTAAAGGCTTAGGATTTGAAGGCGGATGAT 1194  
1355 GATCAAGGA---ATGAGGATGAAGAGTGTGTTCCATGAGGCTTAATAGCGCTG 1411  
1195 GATCAAGGAGTCTTGAAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 1254  
1412 TCACAGGCTGTTTCAGAGCTTGAATATGATGATGATGATGATGATGATGATGATGATGAT 1471  
1255 GCTCAAGGCTGCCCTCAATTTAGATATCTGAGGAGTGTGATGATGATGATGATGATGAT 1314  
1472 TCTCTGGAACATATTGGAATCTTCAAGAACTCTGTTGATTTTTCGCTTGTCTGCTT 1531  
1315 GCTCTAGAGCCATCGGTATACATGAGCAGAGTCTTAAAGATTTCCGCTTGTCTGCTT 1374  
1532 GACCATGAAGAGATTAATGATTTGCTTCAATGAGGCTGAGGCTCTTACTGAGG 1591  
1375 GATAGAGAGCAATATAAAGAGATCACTCCCTTGAAGAGGCTGCTGCTTACTGAGA 1434

QY 1592 GCCTGTGACAACTGAGGAGATTTGCTCTATATCTCAGCGGTGGGGTTGACTGATGTA 1651  
DB 1435 GGTGTCACCAAGCTCCGGAGGTTCCATTTTATGTAAGACCTGGGCTCTCATCATGATGT 1494  
QY 1652 GGCCTTGGTTTACATTTGAGCAATACAGTCCAAATGTGAGATGATGCTGCTGGTTATG 1711  
DB 1495 GGCCTTGGTTTACATTTGAGCAATTTAGTAAGCAATCCGTTTACATTTTCTGGGAATGT 1554  
QY 1712 GGGAGTCTGATGAGGCTTTTGGAGTTCGTAAGGGTGTCTAGTCTTCAGAAAATTT 1771  
DB 1555 GGGAGTCTGAGCAAGGCTTCTCAACTCTCAACAGGATGCCAAGCTTTCGCAAAATG 1614  
QY 1772 GAAATGAGAGGCTTGTATTTTTCAGTGAACGTGCACTTGTCTGGCTGCAACAAATG 1831  
DB 1615 GAGCTGAGGGTGTGT---TCTTCACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1671  
QY 1832 ACTTCTTCTAGTACTTGGGTGCAAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891  
DB 1672 AATCACTGAGATACCTATGGGTGCAAGGATCAAGGCTTCTCCAAATGGCACTGATCTA 1731  
QY 1892 TTGGTAATGGCTGAGCCCTTTTGGAACTTGGATGATGCTTCTAGAAAGGCTGCTACG 1951  
DB 1732 ATGGCAATGGTGGGCTTCTGGAACATTTGGAACATTTGCTGCTGCTGCTGCTGCTGCT 1777  
QY 1952 AATCACTGAGATGAGACTGTAGTGTGAGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1777  
DB 1778 -----ATCAAGATGAAGTTTCCCGGAGCTGAGGACAGATTTCTTGCATATTAT 2011  
QY 2012 TCTTGTGAGGCGAGATGAGATTTTCAGATATCTGTTGCTGCTTTCGCTGCTGCTGCTGCTGCT 2071  
DB 1828 TCATTTGAGGATGAGATGAGATTTTACCTCCTCAGTAAATTTCCCTCTATCTTCTGCTGCT 1887  
QY 2072 TCGGTTGATA 2081  
DB 1888 TGAGTGTATA 1897

RESULT 11  
AK100694

LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J023114118, full insert sequence.

ACCESSION AK100694.1 GI:32985903

VERSION FL1\_CDNA; CAP trapper.

KEYWORDS Oryza sativa (japonica cultivar-group)

SOURCE Oryza sativa (japonica cultivar-group)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

## REFERENCE

1 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team., Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Hotta, I., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group., Ohtsuki, K., Kuroda, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Iida, Y., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Miura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Riken, Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Saito, R., Sasaki, D., Sato, K., Konno, H., Miyazaki, A., Ohtsuki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

22752273

## TITLE

## JOURNAL

## MEDLINE





```
1472 TCTCTGGAACATATTGGAACTCACTTGAAGAACCTCTGTGATTTTCGCTTGGCTTGTGCTT 1531
1898 GCTCTAGAACCCATCGGTACATACAGACAGCTCTTAACGATTTCCGGCTTGTCTGCTT 1957
1532 GACCATGAAGAGAAGAACTAATGATTTGACATTTGCAATGGGGTGGAGGCTCTACTGAGG 1591
1958 GATAGAGAGCAATATATACAGAAATACACCCCTTGACATGGGGTGGCTTACTGAGA 2017
1592 GCTGTGACAACTGAGGAGATTGCTCTATATCTTATATCTCAGCGGTGGCGGTGATGATGA 1651
2018 GCTGACCAAGCTCCGAGGTGGCATTTATGTAAGAGCTGGGGCTCTATCAGATGTT 2077
1652 GCGCTGTGATCATTTGAGCAATACACTCCAAATGTGAGATGATGCTGCTTGGTTATGTG 1711
2078 GCGCTTTGGGTACATCTGCTGATTTAGTAAGACATCCGTTACATGTTACTTGGGAATGT 2137
1712 GGGGAGTCTGATGACAGGGCTTTTGGAGTTTCGCTAAGGGTGTCTAGTCTTCAGAACTT 1771
2138 GGGGAGTCTGACCAAGAGCTGCTTCAACTCTCAACAGGATGCCAAGCTTGCAAAATG 2197
1772 GAATGAGAGGGTGTATTTTTCAGTGAACGTGACTTGTCTGCTGGCTGCAACACAAATG 1831
2198 GAGCTGAGGGGTGT---TTCCTTCACTGACGTCATGCTGCTGCTGTTTACAACTT 2254
1832 ACTTCTCTTAGGTACTTGTGGGTGCAAGGTTATGTTATCTCCATCTGACGCTGATCTT 1891
2255 AATCACTAGATACCTATGTTGGGTGCAAGATACAAAGCATCTCCCAATGGCACTGATCTA 2314
1892 TTGGTAAATGCTGACCCCTTTTGGAACTGATGAGTTCATGCTTCTAGAAAGTGGCTACG 1951
2315 ATGGCAATGCTGGACCCCTTCTGGAACATGAGATTATTGCCCGA----- 2360
1952 AATACCAATCCAGATGAGACTGATGTTGTGACATCCTGCTCATATCTTGTGCATATTAT 2011
2361 -----ATCAAGATGAAGTTTGGCCGAGCGTGAGCAGAGATCTTGTGATATTAC 2410
2012 TCTCTTGCAGGCGAGATCAGATTTTCCAGATACCTGTTGCTGCTTTGACACTGCGACA 2071
2411 TCACTTGCAGAAATGAGATCAGATTACCTTACCTCAGTATTCCTCTATCCTTCGGTT 2470
2072 TGGCTTGTGATA 2081
2471 TGAGTGTGATA 2480

RESULT 12
AF002109/c
LOCUS
DEFINITION
Arabidopsis thaliana chromosome 2 clone T28M21 map C1C10A06,
complete sequence.
ACCESSION
AF002109
VERSION
AF002109.3 GI:20198304
KEYWORDS
HTG.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurooids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 108847)
Rounsley,S.D., Ketchum,K.A., Lin,X., Phillips,C.A., Brandon,R.C.,
Fuhrmann,J.L., White,O., Kerlavage,A.R., Adams,M.D.,
Somerville,C.R. and Venter,J.C.
Unpublished
2 (bases 1 to 108847)
Lin,X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
3 (bases 1 to 108847)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
```

COMMENT  
FEATURES  
source

On Apr 18, 2002 this sequence version replaced gi:6598791.

## Location/Qualifiers

1..108847

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="2"

/map="C1C10A06"

/clone="T28M21"

1..18752

/note="overlap with BAC clone T517

(AC003000:73873..92624)."

complement(2875..2905)

/rpt\_family="AT\_rich"

complement(5687..5709)

/rpt\_family="AT\_rich"

complement(8441..8512)

/gene="At2g39860"

complement(8441..8512)

/gene="At2g39860"

/product="tRNA-Asp"

12204..14036

/gene="At2g39870"

/note="synonym: T28M21.3; supported by cDNA:

gi\_13877742\_gb\_AF370134\_1\_AF370134"

join(12204..12630,12715..13240,13330..13431,13520..14036)

/gene="At2g39870"

join(12358..12630,12715..13240,13330..13431,13520..13611)

/gene="At2g39870"

/codon\_start=1

/product="expressed protein"

/protein\_id="AA095272.1"

/db\_xref="GI:2088640"

/translation="MSRLPPHLNVNLTMTKTSSELTPTRLGPNFEPFDFDPSF  
SGFTSPGDSSTEDDEDFLAGLTERLAPSTORLPSPLEKSEKROVAATSPQ  
STLSGLGSFNSGSRSPILPFPATPSFRDNANDVISAAGEVARLKLGSVEPHL  
PIQTPESLRLQRAIAHAEHQRLIQMWLCSARFKLSNRIPRVNVNBEGLFEN  
PYRVNNPTLWLPQQAAPLAPKPSAGTGVFLPRYPSPAPSDSLKTPVNTAMLPQK  
VKPQNLNFBFTNIVGPRRSQDFECMLARSTVLARQGNFRAVSGGGLPQDMWY"  
complement(13790..13823)  
/rpt\_family="(TA)n"  
14264..14329  
/rpt\_family="(GAAA)n"  
complement(16606..18140)  
/gene="At2g39880"  
/note="synonym: T28M21.4; supported by cDNA:  
gi\_58233308\_gb\_AF175988\_1\_AF175988"  
complement(join(16606..17390,17822..18140))  
/gene="At2g39880"  
complement(join(16606..17390,17822..18140))  
/codon\_start=1  
/product="putative MYB family transcription factor"  
/protein\_id="AA095273.1"  
/db\_xref="GI:2088641"

repeat\_region

repeat\_region

gene

mRNA

CDS

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region



repeat_region	complement (23441..23470) /rpt_family="AT_rich"	1197	ATGCAATGCTAGACGAGGAGTCACTTATGTAATCCAAAGGTGTCCTCAATCTGGAAG	1256
repeat_region	24221..24270 /rpt_family="AT_rich"	42383	ATGCAATATGCTGCTCTCTTTAGTCTTTACATCTATTAAGACTGACAACTCTGTTGTTT	42324
gene	24452..26985 /gene="T28M21.5; supported by full length cDNA: Ceres:39048".24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26985) /gene="At2g39990"	gene	34721..40993 /gene="At2g39930"	40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1317	AGAGCTAAAAAGGCTTAGGATTGAAAGGGCGATGATGATCAAGGATGAGGATGAAG	1376
repeat_region	24635..24667 /rpt_family="GAA)n"	42263	AGCAGTTGAAGCGGCTGAGGATTGAACCGCGTGCAGATGAACAAGGAATGGAAGCGAAG	42204
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1377	AAGTACTGTGCTCCATAGAGGCTTAATAGCTTGTCAACGGGCTGTTTCAAGCTTGAAT	1436
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1437	ACATGGCTGTTTATGCTGATATACAAATGCACTCTCGAATCTTGGAACTCACT	1496
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1497	TGAAGAACTCTGTGATTTTCCCTTGTGCTTGTGCTGACATCAAGAGAGATACTGAT	1556
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1557	TGCCACTTGCATATGAGGCTGCTTACTGAGGGCTGTGCAAGCTGAGGAGATTG	1616
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1617	CTCTATATCTCAGGCGTGGCGGTTGACTGATGTAGGCTTTGGTTTACATTTGACATACA	1676
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1677	GTCCAAATGTGAGATGATGCTGCTTGTGTTATGCGGGAGTCTGATCAGCGGCTTTGG	1736
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1737	AGTTTCGCTAAGGGGCTCTCTAGTCTTCAAGAACTTGAATGAGAGGCTTTTATTTTCA	1796
CDS	Join(24683..24782,25154..25378,25457..25551,25633..25822, 25917..26041,26186..26401,26486..26863) /gene="At2g39990"	CDS	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"	Join(34739..35068,35158..35340,35422..35514,35610..35762, 36241..36332,36678..36819,36895..37002,37419..37504, 37864..37935,38018..38144,38499..38585,38989..39081, 39218..39373,39528..39611,39952..40032,40119..40201, 40384..40501,40581..40993) /gene="At2g39930"
repeat_region	24635..24667 /rpt_family="GAA)n"	1843	AAATCTCAAGAGGCTGTCCAAATCTACAGAACTAGAGTAGAGGTTGT---TGTCTCA	41787

QY 1797 GTGAACGTGCACTGTGTGGCTGCAACAACTTACCTTCTTAGTACTTGTGGTGC 1856  
 Db 41786 GTGAGGAGCAATCGTCGAGCGGTATCAAAATTCCTTCACTGAGATCTGTGGGTAC 41727  
 QY 1857 AAGGTTAAGGTGTATCTCCATCTGACGCTGATCTTTTGGTAATGCTGACCCCTTTTGA 1916  
 Db 41726 AAGGTTACAGAGCATCGATGACGGGACAAGATCTAATGACAGATGGCTAGACCGTACTGGA 41667  
 QY 1917 ACATTGAGTTGATCTTCTAGAAAGGTGGCTACGATACCAATCCAGATGAGACTGTAG 1976  
 Db 41666 ACATGAGCTGATTCATCAAGAAGAGTCCCGAAGTGAATCAACAAGGAGAGATAAGAG 41607  
 QY 1977 TTGTTGAGCATCTGCTCATATCTTGTGCATATTAATCTCTTGACGGGACAGATCAGATT 2036  
 Db 41606 AGATGGAGCATCGGCTCATATATTGGCTTACTACTCTCTGGCTGCCAGAGACAGATT 41547  
 QY 2037 TTCCAGATCTGTTGTCCTTTGGACATGCGACATGCGTTGATCCTAGAGCGGACAGGC 2096  
 Db 41546 GTCCAAACACTGTTAGAGTCTGAAGGAGCCAAATATGATG--ACCCAAAAACAGGTT 41489  
 QY 2097 TGTGTATATATACCAGTTTCTTTTGTGTTT 2126  
 Db 41488 TGTATATAAGATTTTGTCTCGAGTTT 41459  
 RESULT 13  
 AX046779  
 LOCUS 482 bp DNA linear PAT 15-DEC-2000  
 DEFINITION Sequence 5 from Patent WO0068406.  
 ACCESSION AX046779  
 VERSION AX046779.1 GI:11876298  
 KEYWORDS Glycine max (soybean)  
 SOURCE Glycine max  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 REFERENCE 1  
 AUTHORS Caimi, P.G., Fanodu, O.O., Lee, J.M., Miao, G.H. and Maxwell, C.A.  
 TITLE Disease resistance factors  
 JOURNAL Patent: WO 0068406-A 5 16-NOV-2000;  
 E.I. DU PONT DE NEUMOURS AND COMPANY (US)  
 FEATURES  
 Source 1..482  
 /organism="Glycine max"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3847"  
 ORIGIN  
 Query Match 18.0%; Score 412; DB 6; Length 482;  
 Best Local Similarity 97.7%; Pred. No. 1.8e-86;  
 Matches 471; Conservative 0; Mismatches 5; Indels 6; Gaps 5;  
 QY 334 GCGGAAGACACGTTGTGTGTCGACGTGGTCTCGACTGCTGATCCCTTACATCGACGCC 393  
 Db 1 GCGGAAGACACGTTGTGTGTCGACGTGGTCTCGACTGCTGATCCCTTACATCGACGCC 60  
 QY 394 CAAGGACCGGACGCGCGTTTCCAGGTGTGTGCAAGTGTGAGCTGAGCTCGCTCAC 453  
 Db 61 CAAGGACCGGACGCGCGTTTCCAGGTGTGTGCAAGTGTGAGCTGAGCTCGCTCAC 120  
 QY 454 CCGCAAGCAGTCAACATCGCTCTGCTACACACACACCCCGGCTCGCTCGCGCGCG 513  
 Db 121 CCGCAAGCAGTCAACATCGCTCTGCTACACACACACCCCGGCTCGCTCGCGCGCG 180  
 QY 514 CTTCGCCGACCTCGAGTGTCTCAAGCTCAAGGGCAAGCCCGGAGCCCAATGTTCAACT 573  
 Db 181 CTTCGCCGACCTCGAGTGTCTCAAGCTCAAGGGCAAGCCCGGAGCCCAATGTTCAACT 240  
 QY 574 GATACCCGAGTTGGGGCGGACACGTCATCTCCCTGGTCAAGAGATTTCTC-AGTACT 632  
 Db 241 GATACCCGAGATTTGGGGCGGACACGTCATCTCCCTGGTCAAGAGATTTCTCAAGTACT 300

QY 633 TCGATTGCTCAAGAGCCTCCACTTCCGCCGATGATTGTCAA--GGATTCCGATCTTCAG 691  
 Db 301 TCGATTGCTCAAGAGCCTCCACTTCCGCCGATGATTGTCAAAGGGAATTCGATCTTCAG 360  
 QY 692 AATCTCGTGTGACCGCGGTCACTGCTTCAAGCTCTC--AAGCTTGACAAGTGTCTCGG 750  
 Db 361 AATCTCGTGTGACCGCGGTCACTGCTTCAAGCTCTC--AAGCTTGACAAGTGTCTCGG 420  
 QY 751 TTTTACACACCGATGCTCTTTTCCATATC--GTCGCTTTTTCGAA--GAGTTTAAAGAGTCTT 807  
 Db 421 TTTTACACACCGATGCTCTTTTCCATATCGGTCGCTTTTTCGAAAGGTTTAAAGAGTCTT 480  
 QY 808 GT 809  
 Db 481 GT 482

AP003279 134982 bp DNA linear PLN 27-NOV-2003  
 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,  
 PAC clone: P0529E05.  
 ACCESSION AP003279 BA000010  
 VERSION AP003279.2 GI:15718426  
 KEYWORDS  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoidae; Oryzaceae; Oryza.

REFERENCE 1  
 AUTHORS Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,  
 Katayose, Y., Wu, J., Nilmura, Y., Cheng, Z., Nagamura, Y.,  
 Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,  
 Chiden, I., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,  
 Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,  
 Idonuma, A., Iijima, M., Ikeda, M., Ikono, M., Itoh, S., Itoh, T.,  
 Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S.,  
 Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,  
 Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M.,  
 Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,  
 Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,  
 Shomura, A., Song, J., Takasaki, Y., Terasawa, K., Tsuji, K., Waki, K.,  
 Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,  
 Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y.,  
 Yano, M., Jiang, J. and Gojobori, T.  
 The genome sequence and structure of rice chromosome 1  
 Nature 420 (6913), 312-316 (2002)

TITLE  
 JOURNAL MEDLINE  
 PUBMED 12447438  
 REFERENCE 2 (bases 1 to 134982)  
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of  
 Agrobiological Sciences, Rice Genome Research Program, Kamondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,  
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT  
 On Sep 20, 2001 this sequence version replaced gi:13027309.  
 Genes were predicted from the integrated results of the following:  
 GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor  
 (October 1998 version). The genomic sequence was searched against  
 NCBI NonRedundant Protein database, nr  
 (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
 RGP. Protein homologies of the coding regions were searched against  
 NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent  
 the identified cDNA sequences using BLASTN 2.0 with the  
 corresponding DDBJ accession no. and RGP clone ID.  
 A gene with identity or significant homology to a protein is  
 classified based on the protein name to indicate the homology level  
 such as same name, 'putative-' and '-like protein'. A gene without  
 significant homology to any protein but with EST homology (covering

almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.  
 The orientation of the sequence is from T7 to SP6 of the PAC clone. This sequence of P0529E05 clone has an overlap with P0529H11 (DDBJ: AP004072) clone at the position 1 to 19,018 of 5' end. The sequence of this clone starts at the position 146,952 of P0529H11. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

## FEATURES

## source

Location/Qualifiers

1..134382

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="genomic DNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/chromosome="1"

/clone="P0529E05"

complement(788..1069)

/gene="P0529E05.1"

complement(788..1069)

/gene="P0529E05.1"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB84386.1"

/db\_xref="GI:18461189"

/translation="MVASVGGGGRVAAQPRRAGASVATVAGAGGDCGAAERQWR QQDTNQFDEMTRDVRVVELTVKPKFLPVGNHPLVPGFVPVVAIRD"

complement(join(3596..3889,3976..4274,4750..4839,

5288..5456,5562..5684,6335..6361))

/gene="P0529E05.2"

complement(join(3596..3889,3976..4274,4750..4839,

5288..5456,5562..5684,6335..6361))

/gene="P0529E05.2"

/note="contains EST R03279 (R3279)"

/hypothetical protein"

/codon\_start=1

/protein\_id="BAB84387.1"

/db\_xref="GI:18461190"

/translation="MAMIFLLVQEI CMATSNFSEQRNIGNGFTYVKLRDGSIIA VKRATNMVDRHLSERFIEIOTLSKVEHLNLVKFLYHEHERLILVEYVNGSLR EHLDRGLPELESERLNTAIDIVHAYVILGHYTDHPI IHRDITKSNILLTDLRAKY ADGFARLADPNTAEATHTVMVKTAGYVDEPYMRTNQLTDRSDVTSFGVLLVELLTG RRIERGRGRHLLTQWALRKCDDGDAVAMDARRTS AVVAAMEKVMALAECTA PDRAAPAMRCAELVMSIRKDFQHQQRATAAAGARGKHDGTYGPIISLKE"

complement(join(7920..8104,8190..8223,8332..8469,

8893..9039,9136..9180,9283..9646,9758..9984,10424..10843))

/gene="P0529E05.3"

complement(join(7920..8104,8190..8223,8332..8469,

8893..9039,9136..9180,9283..9646,9758..9984,10424..10843))

/gene="P0529E05.3"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB84388.1"

/db\_xref="GI:18461191"

/translation="MAPPGOLLPLASLLPLSAPPFVSGRRRLPTLVLRALPPPTW LPGRSLPAPLPAPRSRPPPPATSLPGASPGGAQAQVLAFTVSERKVA MGLALACNADRVNVAIVLPSQATGYTFAGVVSQSFELGVLSPILGALVDY YGGRVMAYGVAMSLATFLSPANAARSLWLFSTRVLLGAEVGLVPMNNVWRFP PRTSRSAVGIMAGFOLGNTIGLLSPIIMSRAGIFGPFVIFGLFGLVWISAI SGTGEGNAQISAMHLYITRGOKLVKTQSGGERLRKVPFSLSKWPTWALISANAM HSWGYFVLLSWPVPYFKTIVHVNIREAAWFSALPVMVAVLGYVAGVSDRLIQNGTS ILLTRKIMQITGFVGPVALLGLNAKSPVITASLWLTAVGLKSFHSGFLVNLQETIA PQYAGVVLGMSNTAGTFAAILGTVGAGFFVDRMGSFRGLTLLTSLFSSTLFWDIFA TGERVDFDGTG"

11252..11596

/gene="P0529E05.4"

11252..11596

/gene="P0529E05.4"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB84389.1"

/db\_xref="GI:18461192"

/translation="MSSTPAAADTAETDQADLKLIVKAPCGTGRGLITITVTSTQTFADY YFAIKRVKLRVMDLYCGXHSLDPKTKVKFIDDDGRFVRSEQTTEVVGLOQGSISLA IDQGGACIGEN"

complement(join(14141..14355,14614..14942,15117..15432,

15672..15826,15966..16014,16692..18537))

/gene="P0529E05.5"

complement(join(14141..14355,14614..14942,15117..15432,

15672..15826,15966..16014,16692..18537))

/gene="P0529E05.5"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB84390.1"

/db\_xref="GI:18461193"

/translation="MTTLQLDPPVVPARRSAATSCDLHPDETFTGCTACLERLAGL EASAAAAAGPKSTSAIRSLFARFPFGGAGSSVAGAGSLDPLRCKKSCGGRGPD ALAAAAAAGATARADPQRSCDVRGHSTLWSLFGQDRGRVSSSAADIAAPHQ OPPPPPPPTPDDFLDSDI PVMEHDEIMFVVEPVIVDTTSGEITFPNVVAREGKA MKDHMPSSPKPPTKDLKESFWAASVFSKKWKRRQKLKKEAAVSKAAVAAAAM PPEKSPKSFILRRRLRGEAGSELAGRRSCDTPRSLDAGRMSVDDAGFSWDEPR ASWDGLFGAGIGLRAPPLSLKPLPMNSVLETPATIVERSGOQIPVEDADIDP PGGSLQTRDYLDSTRRRLSSSRSPRSPFVTEPAFPAITTIANGKESPLGSEF YHFHAEDLLDRGFSSNLIEDISASLEAALSGPGSAKKPRRWRKASLWGFHRRAS GRSGSPSDIADRSFSEAWPDLRVGNPKMQRCSNLSARSFSSNGSLGSSRRSH VDVGSSARRREHVLERNRARSPPGCRVDMGLRFLYTPMRSGCGGGRARVAAA SETVKGTRRLALIVSMRRLCLPLRLVDLHSSADADTASHTANRPHRCVLSDELL ASIAFVLLAAHGHGRRPTQERTDHSIDRNDGKKHETAGQSSAPASAPAKNHDEP AVASVPRGRKHGRFRAELSTSTRVQMHFARCPPEPGVAKRLATGRNVHVFRR TSSRVPCLGRHMLRQSETGHIRRHGLINAGCCFVSPGVAKRFTCNQSDITKTK OTSETPNFSMRGALILDQMARANCRTHTPTCTAEDMQMHVMSCHMCTIKWSHG STVRNKADMAIVLPAEKVDVETDSTVVRPCHFSVLTKLSIKHLICRSKWI"

complement(join(23661..23843,23923..24021,24116..24188,

24273..24383,24478..24651,24777..25855,25951..27372))

/gene="P0529E05.6"

complement(join(23661..23843,23923..24021,24116..24188,

24273..24383,24478..24651,24777..25855,25951..27372))

/gene="P0529E05.6"

/note="contains EST AU090537 (C12241)"

unknown protein"

/codon\_start=1

/protein\_id="BAB84391.1"

/db\_xref="GI:18461194"

/translation="MADRDEOILYDTIAEGSSVYNEEGNKDPQDLNEEGNVERD AEGNGHMERDVEGQNEEASGSSQVQKARGORGAKKLEGHHITTEVEDRFP SAPAAEAKYVRHSGVVRNDVFNVSFVWRRTARGDHSFVPSDEKMLTMTLTF TPAGTEDVKRMTLKKMAEQFSKGBELIKTILKQTPNFTFKLRDHWDFEYAY KTGEOQQAAMERKENAKKYYHHHLLGGSVSVAMPKMEEMASLLERGLEPATANPE ERSKFWYAHGGTLPNADGSLVFGDQIREAARLTNAVEASSQGTFRDRDELSLA LQTPHPGOTRGKAVIPWKIGFKEDITTVRSQMRSKRDTAKIADLEFRVSSYELSMQ EEVARKVDERMAHRSNNPQPTIPAMVSPSGNRSSCASTGQVSGSMDAMQDQEST CFVDDITQTTCELHLPFKNLSIKVASGNAIPDPSGTYHCRPIPVGYSKVEVELVPG AYVEDLIDYPPGGDGTDLRDTSHAILMHRRYILPRRQASRAPSPAPSPQDPA PSHFAPATPQAPATPPRSPATPPRSPFTTSPQALPAPFASKRPAQAPPAHTR ATKKAIIDAAKNKDPGVDCTQBEDDAYVASVEKRFKPRSPKKIPIDPSVMNFRGM SAPAKAIKLSYVERTLKASGSKVPVQGPQNEIEPLVTGKEMMTKEFTITDYG LTTDQLGVTPIEAKSVKMYEKGKLVKPELLOFLPTOMYKFHLYMENSATGREMI GARIDMDLQGGDILWIPKGYIYQLDALLDLSIMEIORARRRGVDFDTGIDARKV NVAMLDQYQDADENLHLLKAHYIKTIFILLPNTFHHWLLLEDLSACTVNVYDSMD KKESTDFKVFELIDRAWYFRLVGRKRLRRKFKPCAKQKKNKGLNGYVYVCVC HCLADQIITRELDQFIRMNDLTHKEFTITVQEQMLGFIQNEQLDPGGEFYDGNII HRSLSAFLATITTSKS"

complement(30042..34424)

/gene="P0529E05.7"

complement(30042..34424)

/gene="P0529E05.7"

/note="pseudogene, similar to TNP-like transposase element"

/pseudo

complement(join(35905..36225,36435..36510,38277..39001,

42150..42599))

/gene="P0529E05.8"

complement(join(35905..36225,36435..36510,38277..39001,

42150..42599))

/gene="P0529E05.8"

/note="hypothetical protein"

Query Match 16.5%; Score 376.8; DB 8; Length 134982;  
Best Local Similarity 58.8%; Pred. No. 9.7e-78;  
Matches 802; Conservative 0; Mismatches 467; Indels 95; Gaps 5;

QY 798 TAAGAGCTCTGTTTGGAGGAAGCTCAATCTTGAGAAGGACGGAGAAATGGCTACACG 857  
Db 68849 TGAGAACTTTATTTCTGGAGAGTGCCATATTACTGATAGGGTGTGATGGCTTCATG 68908

QY 858 AGCTTGTCTTGAATAATACAGTTCTTGGAGACTCTCAATTTTACTTTGACAGACATTGCTG 917  
Db 68909 AACTTGTGTCAACAACTCTGTTCTGGTGACACTGAACTTCTACATGACTGAACT---CA 68965

QY 918 TTGTGAAGATTGAGGACCTTGAACCTTTAGCTAAATTTGCCCAACTAGTCTGTGA 977  
Db 68966 AAGTGGCGCAGCTGATCTAGAGCTTCTTGAAGAATTTGCAAGTCAATGATTTATTGA 69025

QY 978 AACTTACTGACTGTGAATATCTGATCTTGTGAACCTTTTAAAGCATGCCCTCTGCGCTGG 1037  
Db 69026 AGATGAGTGAGTGATCTTTCAGATCTGATAGTTTCTTCAACAGCAATGGCTGC 69085

QY 1038 AAGAGTTTGTGGAGGCACTTACAGAGGAACAGCA-----AAGATACTCTGCTATAT 1091  
Db 69086 AAGACTTTGTGGAGGAGCATCTTACGAGGTAGGAGAGCTCACCAAGTATGAAAGATTGA 69145

QY 1092 CATTAACGACAAAGTTATGCTGATGGTTTAAACATATATTGGAAGAATGAGTTGCCCA 1151  
Db 69146 AGTTCCCAACCAAGATTATGCTTCTTGGGGCTTACTACATGGGAACAAATGAGATGCTG 69205

QY 1152 TTGTGTTTCATGTTTGCAGCGCTACTTAAATAATTTGGATCTCTCTATGCAATGCTAGACA 1211  
Db 69206 TTATCTTCCCTTTTCGATGAACTCAAGAACTGGACTTGCATATACATTTTCTCAAA 69265

QY 1212 CGGAGGATCATGTTGATGTTAATCCAAAGGTGTCAAATCTGGAAGTCTTGAGACA----- 1267  
Db 69266 CAGAAGATCATTTGTGAGATTATTGCAAAATGTCACCAATCTACTAATTTCTTGAGGTAATAT 69325

QY 1268 ----- 1267

Db 69326 TTCTGTATACAGCAATTGATTAATGTTGTTGAATTAGTATATGCOATATTAACAAG 69385

QY 1268 -----AGGAATGTAATTTGGAGATAGAGGGTTAGAGGTTCTTGCTGCTG 1311  
Db 69386 ATGTCTCTCAACAGGTGAGAACGTTGATAGGAGATAGAGGGCTAGAAAGTTTGTGTATAC 69445

QY 1312 TTGTAAGAGGCTAAAGGCTTAGATTGAAGGGGCGATGATCATCAAGAAATGGAGGA 1371  
Db 69446 ATGCAAGAAGCTACGAAGACTCCGAATTTGAGCGGGTGTATGATGATCCAGGTCTGCAAG 69505

QY 1372 TGAAGAAGGTACTGTGTCCTCATAGAGGCTAATAGCCTTGTACAGGGGCTGTTCAGAGCT 1431  
Db 69506 AGAGCAAGGAGAGTTTCTAGTAGGCTTGACAGCCGTTGCTGTTGTTGCCGTGAAT 69565

QY 1432 TGAATACATGCTGTTTATGTTGCTGATATATACAAATGCTATCTGGAACATATTGGAAC 1491  
Db 69566 GGAGTACATAGTCTGCTATGATCGATATACCAATGCGGCGCTGAGTCTATTGGGAC 69625

QY 1492 TCACCTTGAAGACCTCTGATTTTCGCTTGTGTTGCTTCCACATGAAGAGAGATAAC 1551  
Db 69626 TTTCTGCAAAATCTATACGACTTTTCGCTTGTGCTACTTGAACAGAAAGACAGGTAAAC 69685

QY 1552 TGATTTGCCACTTGCAATGGGGTGAAGGCTCTACTGAGGGGCTGTGCAACAGCTGAGGAG 1611  
Db 69686 AGATCTGCCACTTGCAATGGTGTCTGCTCTGTATAGAAATTTGCAAAAGCTTCGGAG 69745

QY 1612 ATTTGCTCTATATCTCAGCGTGGCGGTTGATGATGAGCCCTTGTTTACATGGAACA 1671  
Db 69746 GTTTGCTCTCTACCTTAGACCAAGAGGGCTTTTCAGATGATGCGCTTAGCTACATCGACA 69805

QY 1672 ATACAGTCCAAATGTGAGATGCTGCTTGGTTGTTATGTTGGGGAGTCTGATGAGGCT 1731  
Db 69806 GTACAGTGGAAATATCCATATATGCTACTGGGCAATGTTGTGTAATCTGACCATGGAT 69865

QY 1732 TTTGGAGTTTGGTAAGGGCTGTCTAGTCTTTCAGAACTTGAATGAGAGGGTGTATT 1791  
Db 69866 GATCCGTTTCGAGTGGGCTGCACCACTTTCAGAAGCTTGAATGAGAAGCTG---CTG 69922

QY 1792 TTTCAAGTAAAGTGCACCTTGTGCTGGCTGCAACAAATGCACTTCTTTAGGTACTTGTG 1851  
Db 69923 CTTTACGCGAGGAGCTTTTGTCCCTGCTTACTGACAGATGCCCTCCCTGAGATACATATG 69982

QY 1852 GGTGCAAGGTTATGGTGTATCTCCATCTGCACTGTATCTTTTGGTAAATGGCTCGACCTT 1911  
Db 69983 GGTGCAAGGATACAGAGCATCTCAAACAGCCCTTGACCTTCTCATGGCAGGCTTT 70042

QY 1912 TTGGAAACATTTGAGTTGATTCCTTCTAGAAGGTGGTACGAATACCAATCCAGATGAGAC 1971  
Db 70043 CTGGAACATCGAGTTTACACCTCCGAGCCCTGAGAGCTTTTAAATCATATGACAGAAGATGG 70102

QY 1972 TGTA-----GTGTTGAGCATCTGCTCATATTTCTTGCATATTTCTTTCAGGCGAGAG 2028  
Db 70103 AGAACCTGTGTGATAGCATGCTCAGGTTCTTGTCTACTATTCCCTTGTCTGGAAGGAG 70162

QY 2029 ATCAGATTTTCCAGATACTGTTGTGCTTTTGGACACTGGCCACAT 2072  
Db 70163 GTCTGATGCCCTCAGTGGGTGATCCCTTGCATCTCGCTGAT 70206

RESULT 15  
AX046791  
LOCUS 1728 bp DNA linear PAT 15-DEC-2000  
DEFINITION Sequence 17 from Patent WO068406.  
ACCESSION AX046791  
VERSION AX046791.1 GI:11876304  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1  
Caimi, P. G., Famodu, O. O., Lee, J. M., Miao, G. H. and Maxwell, C. A.  
Disease resistance factors  
Patent: WO 068406-A 17 16-NOV-2000;  
E.I. DU PONT DE NEMOURS AND COMPANY (US)  
Location/Qualifiers  
FEATURES  
source 1..1728  
/organism="Zea mays"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4577"

ORIGIN

Query Match 16.3%; Score 372.4; DB 6; Length 1728;  
Best Local Similarity 61.0%; Pred. No. 5.3e-77;  
Matches 702; Conservative 0; Mismatches 421; Indels 27; Gaps 5;

QY 807 TGTTTTGGAGGAAGCTCAATTTCTTGAGAGG---ACGGAGATGCTCACAGAGCTTG 863  
Db 34 TTTTCTCTGGAAGAAGCACAAATTTGATGAGAAGAAAATGATGATGATCGTGAGCTTG 93

QY 864 CTTTGAATAATACAGTTCTTGAGACTCTCAATTTTACTTTGACAGACATTGCTGTTGTA 923  
Db 94 CTACAGCAATCTGTTCTTGAGACACTGAATTTCTTTAAACAGATCTCAGGCGATCCC 153

QY 924 AGATTGAGGACCTTGAACTTTTGTAGCTAAATAATGCCCAACTTGTGCTGTGAACTTA 983  
Db 154 CA---GAGTATCTTACCCTTCTGTCGCAACTGTCAACGATTTGAAACTCTGAAGATTA 210

QY 984 CTGACTGTGAATAACTGATCTTGTGACTTCTTTAGCATGCTCTGCGCTGGAAGAGT 1043  
Db 211 GTGAATGTTTCATGCCCACTGCTGCTCAGTTTGTTCGAACTGCAACAACTACAGAGT 270

QY 1044 TTTCTGAGGAGGACCTCAACAGGAGAACAGAAAGATATCTGTCTATA----- 1090  
Db 271 TCGCTGTGTTTCTTTTGAAGAGCAGGCTCACTGTGGCAAGTAGAATATGAGAACT 330

Search completed: April 19, 2004, 22:56:44  
Job time : 8819 secs

24	67.6	3.0	577	3	AAA95064	cdna from
25	49.8	2.2	2000	7	ADA71938	Rice gene
26	48	2.1	660	2	AAV50489	Streptomy
27	48	2.1	7193	2	AAV50431	Streptomy
28	48	2.1	7193	9	AAV50431	Streptomy
29	46.8	2.0	2000	7	ADA71938	Rice gene
30	46.8	2.0	3438	2	AAV39080	S. virido
31	46	2.0	6019	6	ABK31191	Signal tr
32	46	2.0	6019	6	ABL70160	Chemical
33	46	2.0	6019	6	AA61104	Human gen
34	45.4	2.0	17580	6	ABL54312	Chemical
35	45.4	2.0	17580	6	ABN80013	Human che
36	44.8	2.0	1020	3	AAA30718	DNA encod
37	44.8	2.0	1020	9	ADC22700	Human G p
38	44.8	2.0	7327	6	ABL33198	Human imm
39	44.4	1.9	113193	7	AA54645	Streptomy
40	44.2	1.9	6085	2	AA70153	S-longisp
41	44.2	1.9	34094	2	AAZ30163	Complete
42	44.2	1.9	34094	8	ACC78130	Nucleotid
43	44	1.9	2000	7	ADA72854	Rice gene
44	43.6	1.9	732	3	AAA39825	Synthetic
45	43.4	1.9	1020	3	AAA30601	Human G p

ALIGNMENTS

RESULT 1	AAA95063					
ID	AAA95063	standard; cdna; 2288 BP.				
XX	XX	AAA95063;				
AC	AAA95063;					
DT	12-FEB-2001	(first entry)				
DE	DE	cdna from the soybean clone sgs4c.pk003.k23.fis homologous to CO11.				
XX	XX	Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;				
KW	KW	A. thaliana coronatine-induced; CO11; herbicide; EST;				
KW	KW	expressed sequence tag; sgs4c.pk003.k23.fis; 88.				
XX	XX	Glycine max.				
OS	OS					
XX	XX	Location/Qualifiers				
PH	Key	266..2086				
FT	CDS	/*tag= a				
FT	FT	/product= "CO11"				
FT	FT	/partial				
XX	XX	WO200068406-A2.				
PN	PN	16-NOV-2000.				
XX	XX	03-MAY-2000; 2000WO-US011955.				
XX	XX	07-MAY-1999; 99US-0133041P.				
XX	XX	(DUPO ) DU PONT DE NEMOURS & CO E. I.				
PI	PI	Caimi PG, Fawodu OO, Lee J, Miao G, Maxwell CA;				
XX	XX	WPI: 2000-687649/57.				
DR	DR	P-PSDB; AAB23458.				
XX	XX	New nucleic acid sequences encoding new disease resistance factors, and to				
PT	PT	useful for producing plants with increased resistance to pathogens and to				
PT	PT	screen for herbicides.				
XX	XX	Claim 2; Page 59-60; 74pp; English.				
PS	PS	Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA				
XX	XX	libraries were found to be similar to cDNA encoding the Arabidopsis				
CC	CC	thaliana CO11 protein and the Zea mays LLS1 protein. The present sequence				

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:00:14 ; Search time 875 Seconds  
(without alignments)  
11108.431 Million cell updates/sec

Title: US-10-009-791-21

Perfect score: 2288

Sequence: 1 gcacgagccacacgttaca.....acattctcaatttgtgaat 2288

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :	N Geneseq 29Jan04.*
1:	Geneseqn1980s.*
2:	Geneseqn1990s.*
3:	Geneseqn2000s.*
4:	Geneseqn2001as.*
5:	Geneseqn2001bs.*
6:	Geneseqn2002s.*
7:	Geneseqn2003as.*
8:	Geneseqn2003bs.*
9:	Geneseqn2003cs.*
10:	Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2288	100.0	2288	3	AAA95063 cdna from
2	668.4	29.2	2240	3	AAA95062 cdna from
3	412	18.0	482	3	AAN02373 CO11 clon
4	412	18.0	482	3	AAA95055 cdna from
5	372.4	16.3	1728	3	AAA95061 cdna from
6	296.6	13.0	844	3	AAN02372 CO11 cont
7	296.6	13.0	844	3	AAA95054 cdna from
8	233.2	10.2	1074	3	AAA95066 cdna from
9	227.2	9.9	719	6	ABQ65584 Arabidops
10	223.4	9.8	794	3	AAA95056 cdna from
11	223.4	9.8	794	3	AAN02374 CO11 clon
12	193.8	8.5	537	5	AAH87761 Peppermin
13	193.8	8.5	537	5	AAH87873 Peppermin
14	131.6	5.8	1392	8	ADA49036 Wheat gen
15	129	5.6	701	3	AAA95053 cdna from
16	124.4	5.4	283	6	ABL72195 Corn tass
17	114.8	5.0	562	3	AAN02378 CO11 clon
18	114.8	5.0	562	3	AAA95060 cdna from
19	107.8	4.7	486	3	AAA95065 cdna ear-
20	105.2	4.6	266	7	ABX83406 Corn ear-
21	101.6	4.4	2361	3	AAC41105 Arabidops
22	96.2	4.2	1909	3	AAC42586 Arabidops
23	76.4	3.3	609	7	ABZ73247 Rice leaf



is cDNA from the soybean clone sg4c.pK003.k23:fig which is homologous to COIL. The COIL and Lusi proteins are involved in disease resistance. The COIL and Lusi DNAs of the invention may be used to alter the expression of COIL and Lusi protein in cells, particularly to produce transgenic plants with increased systemic resistance to a wide range of pathogens. COIL and Lusi proteins may be used to identify inhibitors of these proteins, which may be useful as herbicides

QY 1981 TGAGCATCCTGCTCATATTTCTTGATATTTATTTCTTTCAGGCGAGAGATCAGATTTTCC 2040  
 Db |||||  
 QY 1981 TGAGCATCCTGCTCATATTTCTTGATATTTATTTCTTTCAGGCGAGAGATCAGATTTTCC 2040  
 Db |||||  
 QY 2041 AGATACTGTTGCTTGGACACTGCGACATGCGTTGATACCTTAGAGGCCAGAGCTGTG 2100  
 Db |||||  
 QY 2041 AGATACTGTTGCTTGGACACTGCGACATGCGTTGATACCTTAGAGGCCAGAGCTGTG 2100  
 Db |||||  
 QY 2101 TATATATACAGATTTTCTTTGTTTCTTCTCCCTTTCATATGCTGTTTCTATGTTCC 2160  
 Db |||||  
 QY 2101 TATATATACAGATTTTCTTTGTTTCTTCTCCCTTTCATATGCTGTTTCTATGTTCC 2160  
 Db |||||  
 QY 2161 TGCTCTATTGTTAGTTCAATTTTATGACAAATAGTCTGTTAAAGGCTGTTTTCATTTG 2220  
 Db |||||  
 QY 2161 TGCTCTATTGTTAGTTCAATTTTATGACAAATAGTCTGTTAAAGGCTGTTTTCATTTG 2220  
 Db |||||  
 QY 2221 AAAATCTGAAGCGTTCCCTTAACGCTATTGGCTCCCTTAAACAACTGAACATTTCAATT 2280  
 Db |||||  
 QY 2221 AAAATCTGAAGCGTTCCCTTAACGCTATTGGCTCCCTTAAACAACTGAACATTTCAATT 2280  
 Db |||||  
 QY 2281 TTGTGAAT 2288  
 Db |||||  
 QY 2281 TTGTGAAT 2288  
 Db |||||

## RESULT 2

AAA95062  
 ID AAA95062 standard; cDNA; 2240 BP.

AC AAA95062;

XX 12-FEB-2001 (first entry)

XX cDNA from the rice clone r10n.pk099.pl4: fis homologous to CO11.

XX Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;

XX A. thaliana coronatine-induced; CO11; herbicide; EST;

XX expressed sequence tag; r10n.pk099.pl4: fis; ss.

XX Oryza sativa.

XX Key Location/Qualifiers

XX CDS 218..2011

XX /\*tag= a

XX /product= "CO11"

XX WO200068406-A2.

XX 16-NOV-2000.

XX 03-MAY-2000; 2000WO-US011956.

XX 07-MAY-1999; 99US-0133041P.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;

XX WPI; 2000-687649/67.

XX P-PSDB; AAB23457.

XX New nucleic acid sequences encoding new disease resistance factors,  
 PT useful for producing plants with increased resistance to pathogens and to  
 PT screen for herbicides.

XX Claim 2; Page 57; 74pp; English.

XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
 CC libraries were found to be similar to cDNA encoding the Arabidopsis  
 CC thaliana CO11 protein and the Zea mays LLS1 protein. The present sequence  
 CC is cDNA from the rice clone r10n.pk099.pl4: fis which is homologous to  
 CC CO11. The CO11 and LLS1 proteins are involved in disease resistance. The  
 CC CO11 and LLS1 DNAs of the invention may be used to alter the expression

CC of CO11 and LLS1 protein in cells, particularly to produce transgenic  
 CC plants with increased systemic resistance to a wide range of pathogens.  
 CC CO11 and LLS1 proteins may be used to identify inhibitors of these  
 CC proteins, which may be useful as herbicides

XX Sequence 2240 BP; 493 A; 586 C; 638 G; 522 T; 0 U; 1 Other;

Query Match 29.2%; Score 668.4; DB 3; Length 2240;

Best Local Similarity 63.4%; Pred. No. 2.5e-177;

Matches 1093; Conservative 0; Mismatches 616; Indels 15; Gaps 4;

QY 353 GACGTGTCCTGACCTGCTCATCCCTTACATCAACGACCCCAAGGACCGGAGCGCTT 412  
 Db |||||  
 QY 290 GAGGAGCGCTGCACCTGCTGCGGTACGTGGACGACCCGCGGACGAGGAGCGGTG 349  
 Db |||||  
 QY 413 TCCAGAGTGTGTCAGCGCTGTACGAGCTCGACTCGCTCACCCGCAAGCAGCTCACCATC 472  
 Db |||||  
 QY 350 TCGTCTGTTGCGCGCTGTGACACCGCATCGAGCGCTTCACCGGACGACGTACCGTG 409  
 Db |||||  
 QY 473 CGCTCTGCTACACACACCCCGCTCGCTCGCGCGCGCTTCCCGCACTCGAGTCG 532  
 Db |||||  
 QY 410 CCTTCTGTAAGCGCGCTCGCCGCGCACCTGCTCGCGCGGTTCCTCGGCTGAGTCG 469  
 Db |||||  
 QY 533 CTCAGGCTCAAGGGCAAGCCCGGAGCCGCAATGTTCAACTTGATACCCGAGATTTGGGCG 592  
 Db |||||  
 QY 470 CTGCGGTCAAGGGGAGCGCGCGCCGCTGTACGGGCTCATCCCGAGGACTGGGCG 529  
 Db |||||  
 QY 593 GGACAGCTCACTCCCTGCTGCTCAAGAGATTTCTCAGTACTTTCGATTGCTCAAGAGCTTC 652  
 Db |||||  
 QY 530 GCTTACGCGCGCCCTCGGTGCGCGGCTCGCGCGCGCTCGCTCGCGCGCGCGCG 649  
 Db |||||  
 QY 713 CACGTGCTTACGCTCTCAAGCTTGACAGGCTCGCGGTTTACCACCGAGTGTCTTTTC 772  
 Db |||||  
 QY 650 CACATGCTGCGAGGAGCTCAAGCTCGACAAAGTGTCTCGCGCTTCTCCACCGACCTCTCCG 709  
 Db |||||  
 QY 773 CATATCGGTGCGCTTTTGAAGAGTTTAAAGAGTCTTTTGGAGGAAAGCTCAATTTCTT 832  
 Db |||||  
 QY 710 CTCGTGCGCGCTCTCTGAGATCACTGAGACATTTATTTCTGAGGAATGCTCAATTTCT 769  
 Db |||||  
 QY 833 GAGAGGACGGAGAAATGCTACACGAGCTTGTGTAATAATACAGTTCTTGAGACTCTC 892  
 Db |||||  
 QY 770 GATAATGTTACTGAATGCTTCCACGACCTTGTCTCAACAATCTCTTCTGAGACATTTG 829  
 Db |||||  
 QY 893 AATTTTACTTGACAGACATTTGCTGTTGTAAGATTTGAGGACCTTTGAACTTTTAGCTAAA 952  
 Db |||||  
 QY 830 AACTTCCACATGACCGAACTCACAGTGTGTTGAGGAGCTTCTCGCAAG 886  
 Db |||||  
 QY 953 AATTGCCCCAACTTAGTGTCTGTAACCTTACTGACTGTGAAATCTGGATCTTTGTGAAC 1012  
 Db |||||  
 QY 887 AAGTGAAGTCACTAATTTCTATTGAAGATCACTGACTGTGACTTTTCAGATTTAATTTGA 946  
 Db |||||  
 QY 1013 TTCTTTAAGCATGCTCTGCGCTGGAAGAGTTTGTGAGGACCTTACAAAGGAGAACCA 1072  
 Db |||||  
 QY 947 TTTTTCGGATGCTGTCATCAATTGAAGAGTTTGGGAGGAGGCGCATTTATTGAGCAAGG 1006  
 Db |||||  
 QY 1073 GA-----AAGATACCTCTGCTATATCAATTACAGCAAGTTATGTCGATTGGTAAACA 1126  
 Db |||||  
 QY 1007 GAGCTCACTAAGTATGGAATGTTAAATTTCCCTTCAAGACTGTGCTTTAGGACTTACG 1066  
 Db |||||  
 QY 1127 TATATTGGAAGAAATGAGTTGCCATTGTTGTTTTCATGTTTTCAGCGCGTACTAAAAAATTG 1186  
 Db |||||  
 QY 1067 TACATGGGACAAACGAGATGCCATTTATCTCCCTTTCTCTGCTTACTCAAGAGCTG 1126  
 Db |||||  
 QY 1187 GATCTCTCTATGCAATGCTAGACACGGAGGATCATTTGTTATTCAAAGGTGTCCA 1246  
 Db |||||  
 QY 1127 GACTTGCACTACATTTTCTCACCCTGAAGATCACTGCCCACTATTGCAAAATGTCCC 1186  
 Db |||||  
 QY 1247 AATCTGGAAGTCTCTGAGACAAAGGATGTAATTTGGAGTAGAGGTTAGAGGTTCTTGGT 1306  
 Db |||||



```

Db 361 AATCTCGCTCGTGCAGCGGTTCACGCTCTCAAGCTTCAAGTGTGCAAGTCTCCGG 420
QY 751 TTTTCCACCGCATGCTCTTTTCCATATC-GGTGCTTTTGCAA--GAGTTTAAAGAGTCTT 807
Db 421 TTTTCAACAACATGCTCTTTTCCATATCGGGTCTGCTTTTGCAAGAAGTTTAAAGTCTT 480
QY 808 GT 809
Db 481 GT 482

RESULT 4
AAA95055
ID AAA95055 standard; cDNA; 482 BP.
XX
AC AAA95055;
XX
DT 12-FEB-2001 (first entry)
XX
DE cDNA from the soybean clone sgs4c.pk003.k23 homologous to COI1.
XX
KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;
KW A. thaliana coronatine-induced; COI1; herbicide; EST;
KW expressed sequence tag; sgs4c.pk003.k23; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT 17..341
FT CDS
FT /product= "COI1"
FT /tag= a
FT /transl_except= (pos:305..308,aa:Xaa)
FT /note= "Xaa= unknown"
FT /partial
XX
XX WO200068406-A2.
XX
PN 16-NOV-2000.
XX
PD 03-MAY-2000; 2000WO-US011956.
XX
PF 07-MAY-1999; 99US-0133041P.
XX
PR (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Cairni PG, Famodu OO, Lee J, Miao G, Maxwell CA;
XX
XX WPI; 2000-687649/67.
XX
DR P-PSDB; AAB23450.
XX
XX New nucleic acid sequences encoding new disease resistance factors,
PT useful for producing plants with increased resistance to pathogens and to
PT screen for herbicides.
XX
XX Claim 2; Page 45; 74pp; English.
XX
XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA
XX libraries were found to be similar to cDNA encoding the Arabidopsis
XX thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence
XX is cDNA from the soybean clone sgs4c.pk003.k23 which is homologous to
XX COI1. The COI1 and LLS1 proteins are involved in disease resistance. The
XX COI1 and LLS1 DNAs of the invention may be used to alter the expression
XX of COI1 and LLS1 protein in cells, particularly to produce transgenic
XX plants with increased systemic resistance to a wide range of pathogens.
XX COI1 and LLS1 proteins may be used to identify inhibitors of these
XX proteins, which may be useful as herbicides
XX
SQ Sequence 482 BP; 96 A; 163 C; 118 G; 105 T; 0 U; 0 Other;

Query Match 18.0%; Score 412; DB 3; Length 482;
Best Local Similarity 97.7%; Pred. No. 2.le-105;
Matches 471; Conservative 0; Mismatches 5; Indels 6; Gaps 5;

```

```

QY 334 GCGAAGACACGTGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 393
Db 1 GCGAAGACACACGTGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 60
QY 394 CAAGGACCGCGACGCGCGCTTTCCAGAGTGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 453
Db 61 CAAGGACCGCGACGCGCGCTTTCCAGAGTGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 120
QY 454 CCGCAAGCAGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 513
Db 121 CCGCAAGCAGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 180
QY 514 CTTCCCGCAGCTCGAGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 573
Db 181 CTTCCCGCAGCTCGAGTGTGTCGACGTGCTCTCGATGCGTCACTCCCTTACATCGAGCCCC 240
QY 574 GATACCCGAGGATTTGGGCGGACACGTCACCTCCCTTGGGTCAAGAGATTTCTCAGTACT 632
Db 241 GATACCCGAGGATTTGGGCGGACACGTCACCTCCCTTGGGTCAAGAGATTTCTCAGTACT 300
QY 633 TCGATTGCTCAAGAGCTTCCATTTCCGCGCATGATTTGTCAA--GGATTCGATTTTCAAG 691
Db 301 TCGATTGCTCAAGAGCTTCCATTTCCGCGCATGATTTGTCAA--GGATTCGATTTTCAAG 360
QY 692 AATCTCGCTCGTGCAGCGGTTCAGTGTCTTCAAGCTCTC--AAGCTTGACAAGTGTCTCGG 750
Db 361 AATCTCGCTCGTGCAGCGGTTCAGTGTCTTCAAGCTCTC--AAGCTTGACAAGTGTCTCGG 420
QY 751 TTTTCAACACCGATGCTCTTTTCCATATC-GGTGCTTTTGCAA--GAGTTTAAAGAGTCTT 807
Db 421 TTTTCAACACCGATGCTCTTTTCCATATC-GGTGCTTTTGCAA--GAGTTTAAAGAGTCTT 480
QY 808 GT 809
Db 481 GT 482

RESULT 5
AAA95061
ID AAA95061 standard; cDNA; 1728 BP.
XX
AC AAA95061;
XX
DT 12-FEB-2001 (first entry)
XX
DE cDNA from the corn clone p0128.cpic134r:fnis homologous to COI1.
XX
KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;
KW A. thaliana coronatine-induced; COI1; herbicide; EST;
KW expressed sequence tag; p0128.cpic134r:fnis; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT 3..1292
FT CDS
FT /tag= a
FT /product= "COI1"
FT /partial
XX
XX WO2000068406-A2.
XX
XX 16-NOV-2000.
XX
XX 03-MAY-2000; 2000WO-US011956.
XX
XX 07-MAY-1999; 99US-0133041P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Cairni PG, Famodu OO, Lee J, Miao G, Maxwell CA;
XX
XX WPI; 2000-687649/67.
XX
XX P-PSDB; AAB23450.
XX
XX New nucleic acid sequences encoding new disease resistance factors,
PT useful for producing plants with increased resistance to pathogens and to
PT screen for herbicides.
XX
XX Claim 2; Page 45; 74pp; English.
XX
XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA
XX libraries were found to be similar to cDNA encoding the Arabidopsis
XX thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence
XX is cDNA from the soybean clone sgs4c.pk003.k23 which is homologous to
XX COI1. The COI1 and LLS1 proteins are involved in disease resistance. The
XX COI1 and LLS1 DNAs of the invention may be used to alter the expression
XX of COI1 and LLS1 protein in cells, particularly to produce transgenic
XX plants with increased systemic resistance to a wide range of pathogens.
XX COI1 and LLS1 proteins may be used to identify inhibitors of these
XX proteins, which may be useful as herbicides
XX
SQ Sequence 482 BP; 96 A; 163 C; 118 G; 105 T; 0 U; 0 Other;

Query Match 18.0%; Score 412; DB 3; Length 482;
Best Local Similarity 97.7%; Pred. No. 2.le-105;
Matches 471; Conservative 0; Mismatches 5; Indels 6; Gaps 5;

```

DR	P-PSDB; AAB23456.	
XX	New nucleic acid sequences encoding new disease resistance factors,	
PT	useful for producing plants with increased resistance to pathogens and to	
PT	screen for herbicides.	
XX	Claim 2; Page 55; 74pp; English.	
XX	Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA	
CC	libraries were found to be similar to cDNA encoding the Arabidopsis	
CC	thaliana COI1 protein and the Zea mays L1S1 protein. The present sequence	
CC	is cDNA from the corn clone p0128.cpic34r:fls which is homologous to	
CC	COI1. The COI1 and L1S1 proteins are involved in disease resistance. The	
CC	COI1 and L1S1 DNAs of the invention may be used to alter the expression	
CC	of COI1 and L1S1 protein in cells, particularly to produce transgenic	
CC	plants with increased systemic resistance to a wide range of pathogens.	
CC	COI1 and L1S1 proteins may be used to identify inhibitors of these	
XX	proteins, which may be useful as herbicides	
SQ	Sequence 1728 BP; 496 A; 341 C; 400 G; 491 T; 0 U; 0 Other;	
	Query Match 16.3%; Score 372.4; DB 3; Length 1728;	
	Best Local Similarity 61.0%; Pred. No. 6.5e-94;	
	Matches 702; Conservative 0; Mismatches 421; Indels 27; Gaps 5;	
QY	807 TGTCTTTGGAGGAAAGCTCAATCTTTGAGAAGG---ACGGAGAAATGGCTACACGAGCTTG 863	
DB		
QY	864 CTTTGAATATACAGTCTTTGAGACTCTCAATTTTACTTTGACAGACATGCTTTGTGA 923	
DB		
QY	924 AGATTGAGACCTTGAACTTTTGTAGCTTAAATAATGGCCCACTAGTCTGTGAAACTTA 983	
DB		
QY	154 CA---GAGTATCTTACCCCTCTTGTGCGCAACTGTCAACGATTGAAACTCTTGAAGATTA 210	
DB		
QY	984 CTGACTGTGAATACAGTCTTTGAGACTCTTGAACCTCTTTAAGCATGCTCTGGCTGGAAGAGT 1043	
DB		
QY	211 GTGAATGTTTCATGCCGATCTGGTCACTGTTGTTCCGAACTGCACAACTACAGAGT 270	
DB		
QY	1044 TTTTGTGGAGCACCTCAACAGAGAACCCAGAAAGATCTCTGCTATA----- 1090	
DB		
QY	271 TCGCTGGTGTCTCTTGAAGAGCAGGGTCAACCTGTGGCAAGTAGAATAATTATGAAACT 330	
DB		
QY	1091 --TCATTACAGCAAGATTATGCTGATGGTTTACATATATTGGAAGAATGAGTTGC 1148	
DB		
QY	331 ACTATTTTCTCTCTCACTGCACCGCTTGATTTGCTCTACATGGGAAACAATGATATGC 390	
DB		
QY	1149 CCATTGTGTTTCATGTTTGCAGCGCTACTAAAAAAATTTGGATCTCTCTATGCAATGCTAG 1208	
DB		
QY	391 AATATCTGTTTCCATATGCTACTGCACTTAAAGATTAGACCTTCAGTTTACATTCCTTT 450	
DB		
QY	1209 ACACGGAGGATCATTTGATGTTTAAATCCAAAGGTGTCAAATCTGGAAGTCTCTTGAGACAA 1268	
DB		
QY	451 CCACAGAGGATCATTTGTGATGATGTTCAAGCTGCTCCAATCTGGAACCTTAGAGGTGA 510	
DB		
QY	1269 GGAATGTTAATTTGAGATGAGGGTTAGAGTCTCTGCTGTTGTTGAAGAGGCTAAAAA 1328	
DB		
QY	511 GGGATGTCATAGGGGATCGTGGATCAAGTTGTTGCACAGACCTGCAAGAAATTCATTA 570	
DB		
QY	1329 GGCTTAGGATTTGAAGGGGGGATGATGATCAAGGA---ATGGAGGATGAAGAAGGTACTG 1385	
DB		
QY	571 GGCTCAGATGAGAGAGGAGATGATGATCAAGGAGGTCTTGAGGATGAACAAGGTAGGA 630	
DB		
QY	1386 TGTCCCATAGAGGCTAATAGCCTTTGTACAGAGGCTGTTTCAAGCTTGATATACATGCGTG 1445	
DB		
QY	631 TTTTACAGGTTGGGTTGATGCTATAGCCCAAGGCTGCCCTGAGTTGACATCTACTGGCGGA 690	
DB		
QY	1446 TTTATGCTCTGATATTAACAATGCTCTCTGGAACATATTGGAACCTCACTTGAAGACC 1505	
DB		
QY	691 TACATGTATACAGATTAACAAATCGACTTTTAGAGGCGAGTTGGTACATGACAGCAAAATC 750	
DB		

QY	1506 TCTGTGATTTTCGGCTTGTGTTGCTTGCACATGAGAGACATAACTGATTTGCCACTTG 1565	
DB		
QY	751 TTAATGACTTCGGCTTGTCTCTCTTGTATAGAGAGCACATATAACCGAATTTGCCACTGG 810	
DB		
QY	1566 ACAATGGGGTGAAGGCTCTACTAGGGGCTGTGACAGCTGAGGAGATTGCTCTATATC 1625	
DB		
QY	811 ACAATGGGGTTCGTGCTTTGCTTTAGAGGTTGCACCAACTACGGAGGTTTGCATTTATG 870	
DB		
QY	1626 TCAGCGTGGGGTGTGATGATGTAGGCTTGGTTACATTTGGACAATACAGTCCAAATG 1685	
DB		
QY	871 TGAGACCTGGGGCCCTATCTGATGTTGCTTGGCTATGTTGGAGAAATTTAGTAAGATA 930	
DB		
QY	1686 TGAGATGATGCTGCTTGGTTATGTGGGGAGTCTGATGAGGGCTTTTGGAGTTGCTTA 1745	
DB		
QY	931 TTCGTTATATGTTGCTTGGTAAATGTTGTAATCTGATAATGAATCATACAAATTTATCAA 990	
DB		
QY	1746 AGGGGTGCTTGTAGTCTTCAGAACTTGAATGAGAGGGTGTATTTTTCAGTGAACGTG 1805	
DB		
QY	991 AAGGCTGCCCAAGCTTGCACAAACTGGAGGTGAGGGTGTGCTCTTT---AGTGAGCATG 1047	
DB		
QY	1806 CACTTGTGCTGTGCTGCAACAAATTAAGTCTTCTTTAGGTACTTGTGGGTGCAAGTTATG 1865	
DB		
QY	1048 CTTTACGCTTGGCTGCACTACAGCTTAAGTCACTGAGGTATCTGTGGGTACCAAGGATTC 1107	
DB		
QY	1866 GTGATCTCCATCTCGAGAGTGTATCTTTTGGTATGCTGACCCCTTTTGGACATTTGAGT 1925	
DB		
QY	1108 GGTCTATCTCCAACTGGAATGATATATTATGGAATGTTAGCCCTTCTGGAACATTTGAGT 1167	
DB		
QY	1926 TGAATTCCTTC 1935	
DB		
QY	1168 ATATTGTCTCC 1177	
DB		
XX	RESULT 6	
XX	AN02372	
ID	AN02372 standard; cDNA; 844 BP.	
XX	AC	
XX	AN02372;	
DT	30-OCT-2001 (first entry)	
XX	COI1 contig of r12.pk0027.h4, r10n.pk099.p14 and r10n.pk0047.c5.	
DE	COI1; rice; herbicide; disease-resistance factor; de.	
XX	Oryza sativa.	
XX	Key	
XX	CDS	
PH	Location/Qualifiers	
FT	9..715	
FT	/*tag= a "COI1"	
FT	/product=	
FT	/partial	
FT	/tranel except= (pos:352..354,aa:Xaa)	
FT	/note= "Xaa= unknown"	
FT	/transl except= (pos:563..664,aa:Arg)	
FT	/note= "This codon has an apparent 1 nucleotide deletion	
FT	which alters the reading frame"	
XX	WO200068406-A2.	
PN	16-NOV-2000.	
XX	03-MAY-2000; 2000WO-US011956.	
XX	07-MAY-1999; 99US-0133041P.	
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX	Calmi PG, Famodu OO, Lee J, Miao G, Maxwell CA;	
XX	WPI; 2000-687649/67.	
DR	P-PSDB; AAP01920.	
XX		

PT New nucleic acid sequences encoding new disease resistance factors,  
 PT useful for producing plants with increased resistance to pathogens and to  
 PT screen for herbicides.

XX Claim 2; Page 44; 74pp; English.

XX An isolated polynucleotide (I) comprising a sequence (Ia) encoding a  
 CC polypeptide (II) of at least 60 amino acids (aa) from any of 18 specified  
 CC sequences, given in the specification, or the complements of (Ia), is  
 CC new. Herbicide. (I) encodes disease-resistance factor polypeptides (II)  
 CC and can be used, in sense or antisense orientation, to alter the  
 CC expression of (II) in cells, particularly to produce transgenic plants  
 CC with increased systemic resistance to a wide range of pathogens. (I) or  
 CC their fragments are also used for gene mapping (e.g. for plant breeding)  
 CC and to identify loss of function mutants. Polypeptides (II) encoded by  
 CC (I) are used to raise specific antibodies (for detection of (II)) and to  
 CC design and/or identify specific inhibitors of (II), potentially useful as  
 CC herbicides. This sequence encodes the rice COI1 protein created from a  
 CC contig of rlr2.pk0027.b4, rlon.pk0099.p14 and rlon.pk0047.c5 as described  
 CC in the method of the invention

XX Sequence 844 BP; 223 A; 171 C; 233 G; 215 T; 0 U; 2 Other;

Query Match 13.0%; Score 296.6; DB 3; Length 844;  
 Best Local Similarity 65.5%; Pred. No. 1e-72;  
 Matches 464; Conservative 0; Mismatches 240; Indels 4; Gaps 2;

QY 1245 CARATCTGGAAGTCTTCCAGACAGGAGTAAATGGAGATAGAGGGTTAGAGGTTCTTG 1304  
 DB 7 CCNACTTACTAGTTCTTCGGTGGAGGATGATTTGGAGATAGAGGATAGGGTTGTTG 66  
 QY 1305 GTCGTTGTTGTAAGAGGCTAAAGGCTTAGGATTTGAAGGGGGGATGATGATCAAGAA 1364  
 DB 67 CAGACACATGCAAGAAGCTCAAGAGCTCAGAGTGGAGGAGATGATGATCCAGGTT 126  
 QY 1365 TGGAGGATGAAGAGGTAAGTGTCCATAGAGGCTAATAGCCCTGTGACAGGCTGTT 1424  
 DB 127 TCCAGAGAAACAGAGGAGGAGTCTCAAGTGGGTTGACAACTGATGCGTAGGATGCC 186  
 QY 1425 CAGAGCTTGAATACATGCTGTTTATGCTGTGATTTACAAATGCAATCTCTGGAACATA 1484  
 DB 187 GTGAATGGAATACATAGTCTGCTAATGCTGTGATATCAAAATGGGCGCTGGAGTCTA 246  
 QY 1485 TTGGAACCTCATTGAAGAACCTCTGTGATTTTTCGCTTGTGTTGATGACCATGAAGAGA 1544  
 DB 247 TTGGGACCTTCTGCAAAATCTTTGCGACTTCCGTTCTGCTACTCGATAGAGAAGAGA 306  
 QY 1545 AGATAACTGATTTGCCACTTGACATGGGGTGGAGGCTCTACTGAGGGCTGTGACAAGC 1604  
 DB 307 GGATAACAGATTGGCCCTTAGACAAATGGTGTCCGTCACCTGCTGANGGGCTGCACGAAC 366  
 QY 1605 TCAGGAGATTGCTCTATATCTCAGCGTGGGGTGTGACTGATGAGGCTTGGTTACA 1664  
 DB 367 TTCGAGGTTTCTCTATCTTGAACACGAGGGGACTTTCAGATACAGGCTTGGCTATA 426  
 QY 1665 TTGGCAATFACAGTCCAAATGTGAGATGGATGCTGCTGGTTATGTTGGGGAGTCTGATG 1724  
 DB 427 TTGGCAGTACAGTGGAAATATCCAAATACATGCTTCTGGGTAATGTTGGGGAACAGATG 486  
 QY 1725 CAGGGCTTTTGAGTTGCTGTAAGGGGTGTCCTAGTCTTCAGAAACTTGAATGAGAGGTT 1784  
 DB 487 ATGGTCTGATCCGGTTTTCATTTGGGGTGTGAGAACCTGCGAAGCTTGA--GCTAAGA 543  
 QY 1785 GTTATTTTTCAGTGAACGTGCACTTGTGCTGCTGCAACACAAATGACTTCTCTTAGT 1844  
 DB 544 GTTGTGCTTCACTGAGCAAGCTTTTAGCCCGGCTATACGGAGTATGCTTTCCTGAGAT 603  
 QY 1845 ACTTGTGGGTGCAAGGTTATGTTGATCTCCATCTGGAAGTATCTTTTGGTAAAGGCTC 1904  
 DB 604 ACGTGTGGGTACAGGGCTACAGGCTTCTAAGACTGCTCAGCATCTCATGCTCATGCC-A 662  
 QY 1905 GACCTTTTGGAAATGATGTTGATTTCTTCTAGAAAGGTGGCTACGA 1952  
 |||||

DB 663 GGCCCTTCTGGAACATAGAGTTTACACCTCCACAGAGACTGGTCAAGA 710

RESULT 7

AAA95054

ID AAA95054 standard; cDNA; 844 BP.

XX AAA95054;

DT 12-FEB-2001 (first entry)

XX cDNA from a rice contig homologous to COI1.

XX Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
 KW A. thaliana coronatine-induced; COI1; herbicide; EST;  
 KW expressed sequence tag; contig; ss.

XX Oryza sativa.

XX Location/Qualifiers

FT Key 9..715

FT CDS /\*tag= a

FT /product= "COI1"

FT /transl\_except= (pos:663..664,aa:Arg)

FT /partial

XX WO200068406-A2.

XX 16-NOV-2000.

XX 03-MAY-2000; 2000WO-US011956.

XX 07-MAY-1999; 99US-0133041P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Calmi PG, Famodu OO, Lee J, Miao G, Maxwell CA;

XX WPI; 2000-687649/67.

DR P-PSDB; AAB23449.

XX New nucleic acid sequences encoding new disease resistance factors,  
 PT useful for producing plants with increased resistance to pathogens and to  
 PT screen for herbicides.

PS Claim 2; Page 44; 74pp; English.

XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
 CC libraries were found to be similar to cDNA encoding the Arabidopsis  
 CC thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence  
 CC is cDNA from a rice contig homologous to COI1. The COI1 and LLS1 proteins  
 CC are involved in disease resistance. The COI1 and LLS1 DNAs of the  
 CC invention may be used to alter the expression of COI1 and LLS1 protein in  
 CC cells, particularly to produce transgenic plants with increased systemic  
 CC resistance to a wide range of pathogens. COI1 and LLS1 proteins may be  
 CC used to identify inhibitors of these proteins, which may be useful as  
 CC herbicides

XX Sequence 844 BP; 223 A; 171 C; 233 G; 215 T; 0 U; 2 Other;

Query Match 13.0%; Score 296.6; DB 3; Length 844;

Best Local Similarity 65.5%; Pred. No. 1e-72;

Matches 464; Conservative 0; Mismatches 240; Indels 4; Gaps 2;

QY 1245 CAAATCTGGAAGTCTTCCAGACAGGATGTAATTTGGAGATAGAGGTTAGAGTTCTTG 1304  
 DB 7 CCNACTTACTAGTTCTTCGGTGGAGGATGTAATTTGGAGATAGAGGTTAGGGTTGTTG 66

QY 1305 GTCGTTGTTGTAAGAGGCTTAAAGGCTTAGGATTTGAAGGGCGGATGATGATCAAGAA 1364

DB 67 CAGACACATGCAAGAAGCTTACAAAGACTCAGAGTTGAGCAGAGATGATGATCCAGGTT 126

QY 1365 TGGAGGATGAAGAAGGTAAGTGTGTGCTCCATAGAGGGCTAATAGCCTTGTTCACAGGGCTGTT 1424



Db 127 TCGAAGAGAACAGAGGAGGTCTCAAGTCGGGTTCACAACTGTAGCCGTAGATGCC 186  
 QY 1425 CAGAGCTTGAATACATGCTGTTTATGCTGTGATTAACAAATCAATCTCTGGAACATA 1484  
 Db 187 GTGAATCGGAATACATGCTGCTATGCTGTGATTAACAAATCGGCTTGGAGTCTA 246  
 QY 1485 TTGGAACCTCACTTGAAGAACCTCTGCTGATTTTTCGGCTTGTGCTTGACCATGAAGAGA 1544  
 Db 247 TTGGGACTTCTCGAAATCTTTGGGACTTCCGCTTGTCTACTCGATAGAGAAGA 306  
 QY 1545 AGATTAACCTGATTTGCCACTTGACAAATGGGTGAGGGCTCTACTGAGGGGCTGTGACAGC 1604  
 Db 307 GGATAACAGATTTGCCCTTAGACAAATGGGTGCTGCTGCTGANGGGCTGCACGAAAC 366  
 QY 1605 TCAGAGAGATTTCTCTATATCTCAGGCGTGGGGGTGACTGATCTAGGCTTGGTTACA 1664  
 Db 367 TTGGGAGGTTTCTCTATATCTGAGACCAAGGGGACTTTCAGATACAGGCTTGGCTATA 426  
 QY 1665 TTGGCAATACAGTCCAAATGTGAGATGATGCTGCTTGGTTATGTTGGGGAGTCTGATG 1724  
 Db 427 TTGGCAGTACAGTGGAAATATCAATACATGCTTCTGGGTAAATGTTGGGAAACAGATG 486  
 QY 1725 CAGGCTTTGAGATTCGCTAGGGGTGCTAGTCTTTCAGAACTTGAATGAGGGGT 1784  
 Db 487 ATGGTCTGATCGGTTTGCATTTGGGGGTGTGAAACCTCGGGAAGCTTGA---GCTAAGGA 543  
 QY 1785 GTTTATTTTTCAGTGAACGTGACCTGCTGTGGCTGCAACACAAATGACTTCTCTAGGT 1844  
 Db 544 GTTGTGCTTTCAGTGAGCAAGTTTATAGCCGCTATACGGGATGCTTCTCTGAGAT 603  
 QY 1845 ACTGTGGGTGAAGTTATGTTATGTTATCTCCATCTGAGCGATGATCTTTTGGTAAATGGCTC 1904  
 Db 604 ACGTGTGGGTACAGGGCTACAAGGCTTCTAAGACTGCTCAAGATCTCATGCTCATGGC-A 662  
 QY 1905 GACCTTTTGGACATGATGTTGATCTTCTTAGAAGGTGGCTACGA 1952  
 Db 663 GGCCCTTCTGGAACATAGAGTTTACACCTCCGAGAACTGTGTACGA 710

## RESULT 8

ID AAA95066 standard; cDNA; 1074 BP.

XX AAA95066;

DT 12-FEB-2001 (first entry)

DE cDNA from the wheat clone wreln.pk0122.d3.fis homologous to COIL.

XX Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
 KW A. thaliana coronatine-induced; COI1; herbicide; EST;  
 KW expressed sequence tag; wreln.pk0122.d3.fis; ss.

OS *Triticum aestivum*.

XX Key Location/Qualifiers  
 FH 2. .667  
 CDS /\*tag= a  
 FT /product= "COIL"  
 FT /partial

XX WO200068406-A2.

PN 16-NOV-2000.

XX 03-MAY-2000; 2000WO-US011956.

XX 07-MAY-1999; 99US-013041P.

XX (DUPO ) DU PONT DE NEMOURS &amp; CO E I.

XX Caiji PG, Famodu OO, Lee J, Miao G, Maxwell CA;

XX WPI; 2000-687649/67.  
 DR P-PSDB; AAB23461.

XX New nucleic acid sequences encoding new disease resistance factors,  
 PT useful for producing plants with increased resistance to pathogens and to  
 PT screen for herbicides.

Claim 2; Page 66; 74pp; English.

XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
 CC libraries were found to be similar to cDNA encoding the Arabidopsis  
 CC thaliana COIL protein and the Zea mays LLS1 protein. The present sequence  
 CC is cDNA from the wheat clone wreln.pk0122.d3.fis which is homologous to  
 CC COIL. The COIL and LLS1 proteins are involved in disease resistance. The  
 CC COIL and LLS1 DNAs of the invention may be used to alter the expression  
 CC of COIL and LLS1 protein in cells, particularly to produce transgenic  
 CC plants with increased systemic resistance to a wide range of pathogens.  
 CC COIL and LLS1 proteins may be used to identify inhibitors of these  
 CC proteins, which may be useful as herbicides

Sequence 1074 BP; 282 A; 232 C; 261 G; 299 T; 0 U; 0 Other;

Query Match 10.2%; Score 233.2; DB 3; Length 1074;

Best Local Similarity 65.7%; Pred. No. 9e-55;

Matches 356; Conservative 0; Mismatches 183; Indels 3; Gaps 1;

QY 1396 AGGGCTAATAGCCTTTGTTCACAGGGCTGTTCAGAGCTTCAATACATATGCTTTATGTGTC 1455

Db 10 AGGATTGATGGCTGTAGCTGAAGGCTGCTCTGATTGGAGTACTGGGCAATACATGTGTC 69

QY 1456 TGATTTACAAATGCAATCTCTGGAACATATTGGAACCTCACTTGRAAGAACTCTGTGATT 1515

Db 70 TGACATTACAAATGCAATCTCTGAGGCTATTGGGCAATTCAGCAAAACCTGGAACGATT 129

QY 1516 TCGCTCTGTTGTTGCTTTCACCATGAAGAGAATACTGATTGCCACTTGCAATGGGGT 1575

Db 130 CGACTTGTCTGCTTGTATAGAGAGTGCATATACTGAACTGCCCTTGCACACGGGT 189

QY 1576 GAGGCTCTACTGAGGGCTGTGACAACTGAGGAGATTGCTCTATATCTCAGGGGTGG 1635

Db 190 TCGGGCTTTGCTGAGAGGTTGCACCAACTCCGAGGTTTGCATTTATGTGAGACCTGG 249

QY 1636 CGGTTGACTGATGAGGCTTGGTTACATTGGCAATACAGTCCAAATGTGAGATGAT 1695

Db 250 AGCTCTATCAGATATGGGCTTTCTTATGTTGGCGAATTTAGCAAGACCGTCCGCTACAT 309

QY 1696 GCTGCTTGTATGTGGGGAGTCTGTGATGCGAGGGCTTTTGGAGTTTCGCTAAGGGGTGTC 1755

Db 310 GTTGTGTTGGAAATCGCGGGGCTCTGATGATGACTGTGTCATTTGCACGAGGATGCC 369

QY 1756 TAGTCTTCAGAACTTGAATGAGAGGGTGTATTATTTTTCAGTGAACGTGCACTGCTGT 1815

Db 370 AAGCTTGCAGAAATTTGGAGCTAA---GGAGTTGCTGCTTTAGTGAACGTGCACTGTCAGT 426

QY 1816 GGCTCAACACAAATGACTTCTCTTAGGTACTTGTGGGTGCAAGGTATTAGGTGATCTCC 1875

Db 427 TGCAGCTTTACAGCTGAAGTCACTCAGATATCTTTGGGTGCGAGGATACAGGCACTCTCC 486

QY 1876 ATCTGAGCGTGTATCTTTTGGTAAATGGCTCGACCCCTTTTGGAAACATTTGATTCCTTC 1935

Db 487 TACTGGCAGCGATCTCTGCGCAATGGTAGCGCCCTTCTGGAACATTTAGTTTATTTGACCC 546

QY 1936 TA 1937

Db 547 AA 548

## RESULT 9

ABQ65584

ID ABQ65584 standard; DNA; 719 BP.

XX

AC ABQ65584;



XX  
DE  
DE  
cor1 clone contig of wrln.pk0122.d3 and wrln.pk0018.fb.

ON 1636 CCGGTTGACTGATGTAGGCCCTTGGTTACATTGGAC-AATACAGTCCAATGTGAGATGGA 1694



PT proteins.  
 XX  
 PS Claim 1; Page 163; 251pp; English.  
 XX  
 CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that  
 correspond to all or part of a mRNA molecule expressed in plant oil gland  
 cells, especially peppermint and plant oil glands that produce terpenoid  
 essential oils and resins. The nucleic acids are useful for genetically  
 mapping a plant genome for genes expressed in plant oil gland cells and  
 to suppress (for example by antisense suppression) or enhance their  
 expression (for example by genetically transforming a plant cell with a  
 replicable expression vector that expresses one or more proteins  
 naturally expressed in plant oil gland cells). The nucleic acids are also  
 useful for recombinant expression of plant oil gland proteins required  
 for terpenoid essential oil and/or resin production in bacterial and/or  
 yeast cells  
 XX  
 SQ Sequence 537 BP; 133 A; 98 C; 138 G; 168 T; 0 U; 0 Other;  
 Query Match 8.5%; Score 193.8; DB 5; Length 537;  
 Best Local Similarity 70.7%; Pred. No. 7.9e-44;  
 Matches 302; Conservative 0; Mismatches 117; Indels 8; Gaps 3;  
 XX  
 CC 1644 CTGATGTAGCCCTTGGTTACATTTGACATACAGTCCAAATGTGAGATGATGCTGCTTG 1703  
 DB 1 CAGACGTGGCCCTTGGTTATATAGGGCAGTACAGATGAAAGTGAGATGATGCTTCG 60  
 QY 1704 GTTATGTGGGGAGTCTGATGACGGCTTTTGGAGTTCGCTAAGGGGTCTCTAGTCTTC 1763  
 DB 61 GGTATGTTGGGGAGTCCGATCAAGGCTTCTAGAAATTTTCCAAAGGGTGCCGAGTCTCC 120  
 QY 1764 AGAACTTGAATGAGAGGGTGTATTTTTCAGTGAACCTGCACCTTGTGTGGCTGCAA 1823  
 DB 121 AGAAGCTTGAATGAGAGGGTGT---TGTATTGAGAGAGACTAGTACAGTCTTC 177  
 QY 1824 CACAATGACTTCTCTTAGGTACTGTGGTGCAAGGTTATGGTGTATCTCATCTGGAC 1883  
 DB 178 TTCAAGTGTGCGCCCTTCGATATTTGTGGTGCAGGATATGTCATCTGGAGATGGTC 237  
 QY 1884 GTGATCTTTTGTGATGCTGCAACCTTTTGGCAACATGAGTGTATCTCTTAGAAGG 1943  
 DB 238 GAGATCTTTTAGCAATGGCCAGACCAATTTGGAATATCGAGTTGATACAGCTCAAGGC 297  
 QY 1944 TGGCTACGAATACCAATCCAGATGAGAC----TGTAGTTGTTGAGCATCTCTCATATT 1999  
 DB 298 ATATTGTTTCATGATGAGAGAGGCAACGATTAGTATGATGTTGAGACCTTGGCATATT 357  
 QY 2000 CTTCATATATCTCTTCAGGGCAGA-GATCAGATTTTCCAGATACATGTTGTGCTTT 2058  
 DB 358 CTTCATATATCTCTTCAGGGCAGAAGGAAATTTGATTTCCCTAGTACTGTTATTCTCT 417  
 QY 2059 GGACACT 2065  
 DB 418 GGATCCT 424  
 RESULT 14  
 ADA49036/c  
 ID ADA49036 standard; DNA; 1392 BP.  
 XX  
 AC ADA49036;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Wheat gene conferring disease resistance in plants.  
 XX  
 KW disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant;  
 wheat.  
 XX  
 OS Triticum aestivum.  
 XX  
 PN WC2003000906-A2.  
 XX

PD 03-JAN-2003.  
 XX  
 PF 21-JUN-2002; 2002WO-IB002453.  
 XX  
 PR 22-JUN-2001; 2001US-03001122.  
 PR 26-SEP-2001; 2001US-03522772.  
 PR 22-MAR-2002; 2002US-03665352.  
 XX  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Glazebrook J, Briggs S, Cooper B, Goff B, Goff SA, Moughamer T;  
 Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;  
 WPI; 2003-184052/18.  
 XX  
 PT New polynucleotide comprising a plant nucleotide sequence having an open  
 reading frame that encodes a polypeptide associated with disease  
 PT resistance, useful for conferring resistance or tolerance to a plant  
 PT pathogen.  
 PS Disclosure; SEQ ID NO 1106; 299pp; English.  
 XX  
 CC The invention relates to a novel isolated polynucleotide comprising a  
 plant nucleotide sequence having an open reading frame that encodes a  
 polypeptide associated with disease resistance or its fragment having  
 substantially the same activity as the full-length polypeptide. The  
 polynucleotide of the invention is useful for conferring resistance or  
 tolerance to a plant pathogen. The present sequence represents a gene  
 conferring disease resistance used in the invention.  
 XX  
 SQ Sequence 1392 BP; 235 A; 441 C; 452 G; 264 T; 0 U; 0 Other;  
 Query Match 5.8%; Score 131.6; DB 8; Length 1392;  
 Best Local Similarity 71.4%; Pred. No. 4.7e-26;  
 Matches 187; Conservative 0; Mismatches 74; Indels 1; Gaps 1;  
 QY 353 GACGTGTCCTCGATCGGTGCATCCCTTACATCGACGACCCCAAGGACCGGACCGGCGCC 412  
 DB 283 GAGGAGGCGCTGCACCTGTGTCTCGCTACGTGGAGACCCCGCGGACCGGAGCGGCC 224  
 QY 413 TCCAGAGTGTGTCGACGCTGTGTACGAGTGCATCGCTCACCAGCAAGCACTCACCATC 472  
 DB 223 TGCTGGCTCCGCGCTGCGACCATCGACGCGCTCAGCGGAGCAAGCACTCACCCTG 164  
 QY 473 GCGCTCTGTACACAC-ACCACCCCGCTCGCTCCGCGCGCTTCCCGCACTCGAGTC 531  
 DB 163 CCCTTCTGTACGCGTGTCTCCCGCGGCTGTCTCGCGCTTCCCGCTCGAGTC 104  
 QY 532 GCTCAAGCTCAAGGCAAGCCCGAGCGCAATGTTCACTTGTATACCGGAGGATGGG 591  
 DB 103 GCTCGGGTCAAGGCAAGCCCGCGCGCATGTACGCGCTCATCCCGAGCACTGGG 44  
 QY 592 CGGACAGTCACTCCCTGGTTC 613  
 DB 43 CGCTACGCGCGCCCTCGGTC 22  
 RESULT 15  
 AAA95053  
 ID AAA95053 standard; cDNA; 701 BP.  
 XX  
 AC AAA95053;  
 XX  
 DT 12-FEB-2001 (first entry)  
 XX  
 DE cDNA from the corn clone p0128.cpici34r homologous to CO11.  
 XX  
 KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
 A. thaliana coronatine-induced; CO11; herbicide; EST;  
 KW expressed sequence tag; p0128.cpici34r; ss.  
 XX  
 OS Zea mays.  
 XX

EH Key Location/Qualifiers  
FT CDS 8..589  
FT /\*tag= a  
FT /product= "COI1"  
FT /partial  
XX  
XX WO200068406-A2.  
XX  
XX 16-NOV-2000.  
XX  
XX 03-MAY-2000; 2000WO-US011956.  
XX  
XX 07-MAY-1999; 99US-0133041P.  
XX  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
XX Caiimi PG, Farnodu OO, Lee J, Miao G, Maxwell CA;  
XX  
XX WPI; 2000-687649/67.  
XX P-PSDB; AAB23448.  
XX  
XX New nucleic acid sequences encoding new disease resistance factors,  
XX PT useful for producing plants with increased resistance to pathogens and to  
XX PT screen for herbicides.  
XX  
XX Claim 2; Page 42-43; 74pp; English.  
XX  
XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
XX libraries were found to be similar to cDNA encoding the Arabidopsis  
XX thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence  
XX is cDNA from the corn clone p0128.cpic134r which is homologous to COI1.  
XX The COI1 and LLS1 proteins are involved in disease resistance. The COI1  
XX and LLS1 DNAs of the invention may be used to alter the expression of  
XX COI1 and LLS1 protein in cells, particularly to produce transgenic plants  
XX with increased systemic resistance to a wide range of pathogens. COI1 and  
XX LLS1 proteins may be used to identify inhibitors of these proteins, which  
XX may be useful as herbicides  
XX  
XX SQ Sequence 701 BP; 193 A; 140 C; 179 G; 180 T; 0 U; 9 Other;  
Query Match 5.6%; Score 129; DB 3; Length 701;  
Best Local Similarity 55.5%; Pred. No. 1.7e-25;  
Matches 318; Conservative 0; Mismatches 234; Indels 21; Gaps 3;  
QY 807 TGTITTTGGAGGAAGCTCAATCTTTGAGAGG---ACGGAGANTGGCTACAGAGCTTG 863  
DB 24 TTTTCTCGGAAGAAAGACAAATGATGAGAAAGAAATGATGAGTGGATCCGTGAGCTTG 83  
QY 864 CTTTGAATAATACAGTTCTTTGAGACTCTCAATTTTACTTGACACAGATTTGCTTTGTGA 923  
DB 84 CTACGAGCAATCTGTTCTTTGAGACACTGAATTTCTTTTACAGATCTCAGGSCATCCC 143  
QY 924 AGATTGAGGACCTTTGAACTTTTAGCTAAATAATGCCCAACTTAGTGTCTGTGAACTTA 983  
DB 144 CA---GAGTATCTTACCCTCTCTGTGGCAACTGTCAACGATTCGAAACTCTGAAGATTA 200  
QY 984 CTGACTGTGAATACCTGGATCTTTGACTCTTTTAGCATGCTCTCGCTGGAAGAGT 1043  
DB 201 GTGAATGTTTCATGCCGATCTGGTCACTTTGTTCCGAACTGCAACACTTCAAGAGT 260  
QY 1044 TTTGTGAGGACCTTACAAACGAGGAACCAAGAGATCTCTGCTATA----- 1090  
DB 261 TCGCTGGTGGTTCTTTTGAAGAGCAGGGTCAACCTGTGGCAAGTAGAAATTTATGAGAACT 320  
QY 1091 ---TCATTAACGAGAAAGTTATGTGANTGGGTTTAACTATATTTGGAAGAAAGATGATTGC 1148  
DB 321 ACTATTTTCTCTCACTGACCGCTTGAGTTTGTCTTACATGGAACAAATGATATGC 380  
QY 1149 CCAATGTGTTTCAATGTTTTCAGCCGCTACTTAAATAATTCGATCTCTCTATCAATGCTAG 1208  
DB 381 AAATACTGTTTCCATATCTACTGCACTTAAAGAGTTAGACCTTCAGTTTACATTCCTTT 440  
QY 1209 ACACGAGGATCATTTGTTATGTTTAAATCCAAAGGTGTCCAAATCTGGAAGTCTCTTGAGACA 1268

DB 441 CCACAGAGGATCATTGNCAGATAGTTCAACGGCTGCTCCAATCTGGAACCTTAGAGGTGA 500  
QY 1269 GGAATGTAATTGGAGATAGAGGGTTAGAGGTTCTTGGTCTGTTGTTAAGAGGCTAAAAA 1328  
DB 501 GGGATGTCATAGGGGATCGGGGACTACAAGNTGGTGCACAGACCTGCAAGAAATTCGATA 560  
QY 1329 GGCTTAGGATTGAAAGGGGCGGATGATGATCAAG 1361  
DB 561 GGCTCAGAGTAGAGAGGAGATGATGATCNAG 593  
Search completed: April 19, 2004, 18:52:44  
Job time : 880 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:30:24 ; Search time 164 Seconds

(without alignments)  
7742.242 Million cell updates/sec

Title: US-10-009-791-21

Perfect score: 2288

Sequence: 1 gcacggccacacgttaca.....acattctcaattgtgaat 2288

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105.2	4.6	266	4	US-09-313-294A-1866
2	54	2.4	7218	1	US-08-232-463-14
3	49.6	2.2	7218	1	US-08-232-463-14
4	46.8	2.0	735	3	US-09-003-287-7
5	44.8	2.0	1020	4	US-09-170-496D-181
6	44.2	1.9	6085	3	US-09-023-603-4
7	44.2	1.9	34094	4	US-09-292-034-1
8	43.4	1.9	1020	4	US-09-170-496D-31
9	43.4	1.9	1900	4	US-09-016-434-1484
10	43.4	1.9	1901	1	US-08-153-848-43
11	43.4	1.9	1901	3	US-09-299-843A-43
12	43.4	1.9	1901	4	US-09-088-337B-43
13	43.4	1.9	1901	5	PCT-US93-11153-43
14	43.4	1.9	2453	5	PCT-US95-07180-1
15	43.2	1.9	51952	3	US-08-947-823-1
16	42.6	1.9	609	4	US-09-252-991A-15483
17	42.6	1.9	1083	4	US-09-252-991A-15422
18	42.6	1.9	1461	4	US-09-252-991A-15258
19	42.6	1.9	1929	4	US-09-252-991A-15253
20	42.2	1.8	1155	2	US-08-387-942C-7
21	42	1.8	1683	4	US-09-252-991A-8337
22	42	1.8	1833	4	US-09-252-991A-8371
23	42	1.8	1380	4	US-09-252-991A-8507
24	41.8	1.8	44377	2	US-08-804-227C-7
25	41.8	1.8	44377	2	US-08-804-198-1
26	40.8	1.8	1926	4	US-09-249-585A-2
27	40.8	1.8	1926	4	US-09-410-399-3

C 28	40.8	1.8	2580	3	US-09-050-863-2	Sequence 2, Appli
C 29	40.8	1.8	2580	4	US-09-359-081-2	Sequence 2, Appli
C 30	40.8	1.8	5452	2	US-09-130-114-1	Sequence 1, Appli
C 31	40.8	1.8	8705	4	US-09-647-344A-14	Sequence 14, Appli
C 32	40.8	1.8	9600	3	US-08-910-647-1	Sequence 1, Appli
C 33	40.8	1.8	9600	4	US-09-620-925-1	Sequence 1, Appli
C 34	40.8	1.8	10596	1	US-07-884-811-15	Sequence 15, Appli
C 35	40.8	1.8	10596	1	US-07-885-971-15	Sequence 15, Appli
C 36	40.8	1.8	10596	1	US-08-087-783A-15	Sequence 15, Appli
C 37	40.8	1.8	10596	2	US-08-194-088B-15	Sequence 15, Appli
C 38	40.8	1.8	10596	5	PCT-US93-04648-15	Sequence 15, Appli
C 39	40.8	1.8	16080	4	US-09-724-566A-48	Sequence 48, Appli
C 40	40.8	1.7	1151	1	US-07-704-288C-2	Sequence 2, Appli
C 41	39.6	1.7	1151	1	US-08-093-372-1	Sequence 1, Appli
C 42	39.6	1.7	1151	1	US-08-379-259-2	Sequence 2, Appli
C 43	39.6	1.7	1151	1	US-08-482-273-49	Sequence 49, Appli
C 44	39.4	1.7	639	4	US-09-410-551B-1	Sequence 1, Appli
C 45	39.2	1.7	77536	4		

#### ALIGNMENTS

RESULT 1  
US-09-313-294A-1866  
; Sequence 1866, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 1866  
; LENGTH: 266  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700551651H1  
US-09-313-294A-1866

Query Match	4.6%	Score 105.2;	DB 4;	Length 266;
Best Local Similarity	64.7%	Pred. No. 3.3e-20;		
Matches 172;	Conservative 0;	Mismatches 93;	Indels 1;	Gaps 1;
QY	1368	AGGAATGTAATTGGAGATAGAGGGTTAGAGGTTCTTGGTTCGTTGTTGTAAGAGGCTAAAA	1327	
Db	2	AGGAATGTAATTGGAGATAGAGGTTAGAGGTTCTTGGGATACGTGCAAGAGCTCCAA	61	
QY	1328	AGGCTTAGGATTGAAAGAGGGCGCATGATCATCAGGATGGAGGATGAGAGGTTACTGTG	1387	
Db	62	AGGCTCAGATAGACGAGGAGATGATGAGGAGGTGTGCAAGAG-ACAGGGAGGGTTC	120	
QY	1388	TCCATAGAGGCTTAATAGCCCTTGTACAGGCGTGTTCAGAGCTTGAATACATGCTGT	1447	
Db	121	TCTCAAGTGGGCTTGACCGGCTATAGCCGTATAGCCGTAGTTCGCTGAGCTGGAATATATAGTGC	180	
QY	1448	TATGTGCTGATATTACAAATGTCATCTCTTGGAAATATTTGGAACCTACTTGAAGAACCTC	1507	
Db	181	TATGTGCTGATATAACCAATGGGGCTTGGAACTATCTGCGGACATCTGCAAAAACTA	240	
QY	1508	TGTGATTTTCGCTTGTTGCTTGA	1533	
Db	241	TACGACTTCGGCTTGTCTTACTTGA	266	

RESULT 2  
US-08-232-463-14













RESULT 15  
US-08-947-823-1/c  
; Sequence 1, Application US/08947823  
; Patent No. 6114605  
; GENERAL INFORMATION:  
; APPLICANT: Williamson, Valerie M.  
; APPLICANT: Kaloshani, Isqouhi  
; APPLICANT: Yaqoobi, Jafar

Tue Apr 20 10:32:12 2004

APPLICANT: Bodeau, John  
 APPLICANT: Milligan, Stephen  
 TITLE OF INVENTION: Procedures and Materials for Conferring  
 TITLE OF INVENTION: Pest Resistance in Plants  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/947,823  
 FILING DATE: 09-OCT-1997  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US97/18802  
 FILING DATE: 09-OCT-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/028,191  
 FILING DATE: 10-OCT-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bastian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-070210US  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 51952 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-947-823-1

Search completed: April 19, 2004, 22:59:27  
 Job time : 166 secs

Query Match 1.9%; Score 43.2; DB 3; Length 51952;  
 Best Local Similarity 46.5%; Pred. No. 0.36; Indels 21; Gaps 3;  
 Matches 264; Conservative 0; Mismatches 283;

50098 -----CTGGTCTCCTTACAGATCGTGTGTTGAGTACATCGGGGTCCATGCTAAGAAGT 50045  
 1686 TGAGATGGATGCTGCTTGGTGTATGTTGGGGAGTCTGATGAGGGCTTTTGGAGTTGCTTA 1745  
 50044 TAGAGATGCTTTTCCCTTAGCTTTTGGGGAGATAGTATATAGGCTTCTCTATGTTCTCT 49985  
 1746 AGGGGTGTCCTAGTCTTTCAGAACTTGAATGAGAGGGTGTATTTTTCAGTGAAAGCTG 1805  
 49984 CTGGTTTTCAGAGGCTTCCGTAAGTTGGAGACTAGAGACTGCCCTTTTGGCGATGAGGCTC 49925  
 1806 CACTTGTGCTGCTGCAACACACATTGAC 1833  
 49924 TGTGGCTAATGCTGCAAGCTGGAGAC 49897

QY 1272 ATGTATTCGAGATAGAGGTTTAGAGTTCTTGGTGGTTTGTGTAAGAGGCTAAAGAGGC 1331  
 DB 50449 ATTATATTGAAGATAGCGTCTTGAGGATTCGCAACACTTGTGAAGAACTTCAAGAGC 50390  
 QY 1332 TTAGGATTTGAAAGGGCGGATGATCAAGGAATGGAGGATGAAGAGTACTGTGTCCC 1391  
 DB 50389 TTAGGCT-----GTTTCCTTTTGATCTATTGTCTCCAGGACCTAATGTATCCTTGACAG 50336  
 QY 1392 ATAGAGGCTAATAGCTTGTACAGGGCTTTCAGAGCTTGAATACATGCGCTTTATG 1451  
 DB 50335 AGCAGGCTTGTAGTGTCTCAATGGGCTGCGCTAAGCTTTCAGTTCAGTTTATCTTCT 50276  
 QY 1452 TGTCTCATATTACAAATGCATCTCTGAAACATATTGGAACCTAC-----TTGAAGAACC 1505  
 DB 50275 GCGGCCAATGCAAAATGAGCGCTTAGTTAGTATTGCNAGGACCGTTCTAACATGATCC 50216  
 QY 1506 TCTGTGATTTTCGCTTGTGTTGCTTGACCATGAAGAGAGATACTGATTCGCACTTG 1565  
 DB 50215 GATTTGTTGTTGTTATTCGAGCCTCAAACTCTGACTACTTAACTCTTGAACCACTTG 50156  
 QY 1566 ACAATGGGGTGAAGGCTCTACTGAGGGCTGTGACAGCTGAGGAGATTGCTCTATATC 1625  
 DB 50155 ATACTGTTTTGGGGCCATTTGTGCAACACTGCAAGGAAGTGCAGCGACTTCTCTTT--- 50099  
 QY 1626 TCAGGCGTGGCGGTTTGCATCTGATGTAGGCTTGGTTACATTGGACAATACAGTCCAAATG 1685

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 20:29:51 ; Search time 1487 Seconds  
(without alignments)  
6884.896 Million cell updates/sec

Title: US-10-009-791-21

Perfect score: 2288

Sequence: 1 gcacaggccacagcttaca.....acatttcatttgat 2288

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/FCI\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/FCIUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2054	89.8	2305	13	US-10-424-599-141824
2	1999	87.4	2039	13	US-10-425-114-11061
3	1982	86.6	2011	13	US-10-425-114-10724
4	609.8	26.7	2178	16	US-10-310-154-152
5	561.2	24.5	2419	16	US-10-310-154-151
6	409.8	17.9	1503	13	US-10-425-114-32228
7	403.8	17.6	1325	13	US-10-425-114-8867
8	394.8	17.3	1469	13	US-10-425-114-33936
9	323	14.1	669	13	US-10-424-599-125710
10	311.6	13.6	744	13	US-10-424-599-125711
11	290.4	12.7	829	13	US-10-424-599-317
12	279.8	12.2	428	13	US-10-424-599-44612
13	227.2	9.9	719	9	US-09-770-149-161
14	220.2	9.6	449	13	US-10-424-599-21331

15	218.6	9.6	789	13	US-10-424-599-21332	Sequence 21332, A
16	199.2	8.7	477	13	US-10-424-599-27395	Sequence 27395, A
17	193.8	8.5	212	13	US-10-424-599-86569	Sequence 86569, A
18	185.8	8.1	350	13	US-10-424-599-38104	Sequence 38104, A
19	179.6	7.8	447	13	US-10-424-599-79984	Sequence 79984, A
20	166.4	7.3	417	13	US-10-424-599-39506	Sequence 39506, A
21	160.2	7.0	361	13	US-10-424-599-32267	Sequence 32267, A
22	125.4	5.5	1142	13	US-10-424-599-97745	Sequence 97745, A
23	124.4	5.4	283	9	US-09-294-0938-1569	Sequence 1569, Ap
24	112.2	4.9	521	13	US-10-424-599-28814	Sequence 28814, A
25	105.8	4.6	712	13	US-10-424-599-5625	Sequence 5625, Ap
26	105.6	4.6	711	13	US-10-424-599-34663	Sequence 34663, A
27	99.6	4.4	2793	13	US-10-424-599-102884	Sequence 102884, A
28	97.2	4.2	2465	13	US-10-425-114-26663	Sequence 26663, A
29	97	4.2	37	9	US-09-969-373-1014	Sequence 1014, Ap
30	96	4.2	2345	16	US-10-310-154-287	Sequence 287, App
31	86.2	3.8	682	13	US-10-424-599-93112	Sequence 93112, A
32	82.8	3.6	274	9	US-09-923-876-1583	Sequence 1583, Ap
33	82.8	3.6	274	11	US-09-923-876-1583	Sequence 1583, Ap
34	81.6	3.6	3433	13	US-10-424-599-91753	Sequence 91753, A
35	75.6	3.3	2693	13	US-10-424-599-69165	Sequence 69165, A
36	70	3.1	269	13	US-10-424-599-130155	Sequence 130155, A
37	68.6	3.0	1236	13	US-10-424-599-34657	Sequence 34657, A
38	67	2.9	1808	13	US-10-424-599-4935	Sequence 4935, Ap
39	65	2.8	302	13	US-10-424-599-83717	Sequence 83717, A
40	58.2	2.5	254	13	US-10-424-599-45283	Sequence 45283, A
41	52.6	2.3	2436	15	US-10-156-761-7456	Sequence 7456, Ap
42	52.6	2.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
43	52.4	2.3	3147	13	US-10-424-599-4932	Sequence 4932, Ap
44	51	2.2	343	13	US-10-424-599-65547	Sequence 65547, A
45	50.6	2.2	327	13	US-10-424-599-137620	Sequence 137620, A

ALIGNMENTS

RESULT 1  
US-10-424-599-141824  
; Sequence 141824, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 141824  
; LENGTH: 2305  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_99079C.1  
US-10-424-599-141824

Query Match	89.8%	Score 2054;	DB 13;	Length 2305;
Best Local Similarity	99.5%	Pred. No. 0;		
Matches 2060;	Conservative	0;	Mismatches 10;	Indels 0;
Gaps	0;			
Qy	219	TTTCTCCTCATCATCCCAATACCATTCCTTTTGGCCACTTGACCAAAACCTCTG	278	
Db	191	TTCTTCTCCTCATCATCATCAACCCATTCCTTTTGGCCACTTGACCAAAACCTCTG	250	
Qy	279	CACCTTTTCTTTTTCACCTCTCGATCCATATATGACGAGGACGAAACGTCGCGA	338	
Db	251	CACCTTTTCTTTTTCACCTCTCGATCCATATATGACGAGGACGAAACGTCGCGA	310	
Qy	339	AGACACGTGTGGTCGAGTGGTCTCTCGACTCGCTTACATCGACGCCCAAGG	398	



```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 11061
; LENGTH: 2039
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 70
US-10-425-114-11061

```

Query Match 87.4%; Score 1999; DB 13; Length 2039;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2010. Conservative 0; Mismatches 0; Indels 1;

Qy	278	GCACCTTTCTTTTCTCACTCTCAGTCTCCGATCCCAATATGACCGAGGACGGAACTGCGG	337
Db	1	GCACCTTTTCTTTTCTCACTCTCAGTCTCCGATCCCAATATGACCGAGGACGGAACTGCGG	60
Qy	338	AAGACAGTGTGCTCGACGTGGTCTCTCGACTCGTCCGTATCCCTTATACATCGACGACCCCAAG	397
Db	61	AAGACAGTGTGCTCGACGTGGTCTCTCGACTCGTCCGTATCCCTTATACATCGACGACCCCAAG	120
Qy	398	GACCGGACGCCGTTTCCAGGTGTGTGACAGCGTGTGTAAGAGTGTGACTCGACTCGCTCAACCGC	457
Db	121	GACCGGACGCCGTTTCCAGGTGTGTGACAGCGTGTGTAAGAGTGTGACTCGACTCGCTCAACCGC	180
Qy	458	ARGACGTACCAATCCGGCTCTGCTACACACACCCCGCTCGCTCCGCGCGCGCTTC	517
Db	181	ARGACGTACCAATCCGGCTCTGCTACACACACCCCGCTCGCTCCGCGCGCGCTTC	240
Qy	518	CCGCACTCTGAGTCGCTCAAGCTCAAGGGCAAGCCCGAGCGCGCAATGTTCAATTGTATA	577
Db	241	CCGCACTCTGAGTCGCTCAAGCTCAAGGGCAAGCCCGAGCGCGCAATGTTCAATTGTATA	300
Qy	578	CCGAGGATTTGGGGCGGACAGCTCACTCCCTGGGTCAAGAGATTTCTCAGTACTTTCGAT	637
Db	301	CCGAGGATTTGGGGCGGACAGCTCACTCCCTGGGTCAAGAGATTTCTCAGTACTTTCGAT	360
Qy	638	TGCCTCAAGAGCTTCCACTTTCGCGCGCATGATTTGTTCAAGGATTTCCGATCTTCAGAACTCTC	697
Db	361	TGCCTCAAGAGCTTCCACTTTCGCGCGCATGATTTGTTCAAGGATTTCCGATCTTCAGAACTCTC	420
Qy	698	GCTCGTGACCGCGGTCAAGTGTCTCAAGCTTCAAGCTTGAAGTGTCTCCGGTTCAC	757
Db	421	GCTCGTGACCGCGGTCAAGTGTCTCAAGCTTGAAGTGTCTCCGGTTCAC	480
Qy	758	ACCGATGGTCTTTTCCATATACGGTCGGCTTTTGAAAGAGTTTAAAGTCTTGTGTTTGGAG	817
Db	481	ACCGATGGTCTTTTCCATATACGGTCGGCTTTTGAAAGAGTTTAAAGTCTTGTGTTTGGAG	540
Qy	818	GAAGAGCTCAATCTTTGAGAGGACGGAGATTGGCTACAACGAGCTTCTTTGAAATAATA	877
Db	541	GAAGAGCTCAATCTTTGAGAGGACGGAGATTGGCTACAACGAGCTTCTTTGAAATAATA	600
Qy	878	GTTCTTGAGACTCTCAATTTTATCTTGACAGACATTCGCTGTGTGAGATTTGAGGACCTT	937
Db	601	GTTCTTGAGACTCTCAATTTTATCTTGACAGACATTCGCTGTGTGAGATTTGAGGACCTT	660
Qy	938	GAATCTTTTAGCTTAAAAATGCCCCAACTTAGTGTCTGTGAAACTTACTGACTGTGAAATA	997
Db	661	GAATCTTTTAGCTTAAAAATGCCCCAACTTAGTGTCTGTGAAACTTACTGACTGTGAAATA	720
Qy	998	CTGGATCTTTGTGAACTTTTAAAGCATGCTCTGGCTGGAGAGATTTTGTGGAGGCACC	1057
Db	721	CTGGATCTTTGTGAACTTTTAAAGCATGCTCTGGCTGGAGAGATTTTGTGGAGGCACC	780
Qy	1058	TACAACGAGGACCAAGAGATPACTCTGCTATATCATTTACCAGCAAAAGTTATGTCCGATTG	1117
Db	781	TACAACGAGGACCAAGAGATPACTCTGCTATATCATTTACCAGCAAAAGTTATGTCCGATTG	840
Qy	1118	GGTTTAAACATATATTGGAAAGATGAGTTGCCCATTTGTTTCATGTTTGACGCGTACTA	1177
Db	841	GGTTTAAACATATATTGGAAAGATGAGTTGCCCATTTGTTTCATGTTTGACGCGTACTA	899

QY	1178	AAAAATGGGATCTCTCTCTATGCAATGCTAGACACGGAGGATCATTTGATGTATATATCCAA	1237
DB	900	AAAAAATTTGGATCTCTCTCTATGCAATGCTAGACACGGAGGATCATTTGATGTATATCCAA	959
QY	1238	AGGTGTCAAAATCTGGAAGTCCTTCAGACACAGGAATGTAAATGGAGATAGAGGGTTAGAG	1297
DB	960	AGGTGTCAAAATCTGGAAGTCCTTCAGACACAGGAATGTAAATGGAGATAGAGGGTTAGAG	1019
QY	1298	GTTCTTTGGTCGTTGTTGAAGAGGCTAAAAGGCTTAGGATTTGAAGGGGGCGATGATGAT	1357
DB	1020	GTTCTTTGGTCGTTGTTGAAGAGGCTAAAAGGCTTAGGATTTGAAGGGGGCGATGATGAT	1079
QY	1358	CAAGGAATCGAGGATGAAGAAAGGTACTGTGTGCCATAGAGGGCTAATAGCCTTTGTCCACAG	1417
DB	1080	CAAGGAATCGAGGATGAAGAAAGGTACTGTGTGCCATAGAGGGCTAATAGCCTTTGTCCACAG	1139
QY	1418	GGCTGTTCCAGAGCTTGAATATACATGGCTGTTTATGTCTGTGATATTTACAAATGCATCTCTG	1477
DB	1140	GGCTGTTCCAGAGCTTGAATATACATGGCTGTTTATGTCTGTGATATTTACAAATGCATCTCTG	1199
QY	1478	GAACATATTTGGAACTCACATTTGAAGAACCTCTGTGATTTTCGGCTTGCTGTGCTTGACCAAT	1537
DB	1200	GAACATATTTGGAACTCACATTTGAAGAACCTCTGTGATTTTCGGCTTGCTGTGCTTGACCAAT	1259
QY	1538	GAAGAGAAGATACTGATTTGCCACTTTGACAAATGGGGTGAGGGCTCTACTGAGGGGCTGT	1597
DB	1260	GAAGAGAAGATACTGATTTGCCACTTTGACAAATGGGGTGAGGGCTCTACTGAGGGGCTGT	1319
QY	1598	GACAGCTGAGGAGATTTGCTCTATATCTACGGCTGGGGTGAGCTGATGTAGGSCCTT	1657
DB	1320	GACAGCTGAGGAGATTTGCTCTATATCTACGGCTGGGGTGAGCTGATGTAGGSCCTT	1379
QY	1658	GGTTCATATTTGGACAAATACAGTCCAAATGTGAGATGCATGCTGCTGTTGTTATGGGGGAG	1717
DB	1380	GGTTACATTTGGACAAATACAGTCCAAATGTGAGATGCATGCTGCTGTTGTTATGGGGGAG	1439
QY	1718	TCTGATGCAGGGCTTTTGGAGTTTCGCTTAAGGGTGCTCTAGTCTTCAGAAACTTTGAAATG	1777
DB	1440	TCTGATGCAGGGCTTTTGGAGTTTCGCTTAAGGGTGCTCTAGTCTTCAGAAACTTTGAAATG	1499
QY	1778	AGAGGTGTTTATTTTTCAGTGAACGTGCACTTGTCTGTGGCTGCAACACAAATTTGACTTCT	1837
DB	1500	AGAGGTGTTTATTTTTCAGTGAACGTGCACTTGTCTGTGGCTGCAACACAAATTTGACTTCT	1559
QY	1838	CTTAGGTACTTGTGGGTGCAAGGTTATGTGTATCTCCAFCTGGAACGTGATCTTTTGGGTA	1897
DB	1560	CTTAGGTACTTGTGGGTGCAAGGTTATGTGTATCTCCAFCTGGAACGTGATCTTTTGGGTA	1619
QY	1898	ATGGCTCGAACCTTTTGGGAAACATTTGAGTTGATTCCTCTAGAAAGGTGGCTACGAATACC	1957
DB	1620	ATGGCTCGAACCTTTTGGGAAACATTTGAGTTGATTCCTCTAGAAAGGTGGCTACGAATACC	1679
QY	1958	AATCCAGATGAGACTGTAGTTGTTTGAGCATCTGCTCATATCTTTCGATATATATCTCTT	2017
DB	1680	AATCCAGATGAGACTGTAGTTGTTTGAGCATCTGCTCATATCTTTCGATATATATCTCTT	1739
QY	2018	GCAGGCGAGAGATCAGATTTTCCAGATCTGTTGTGCTTTTGACACTCGCCACATGCGTT	2077
DB	1740	GCAGGCGAGAGATCAGATTTTCCAGATCTGTTGTGCTTTTGACACTCGCCACATGCGTT	1799
QY	2078	GATACCTTAGAGGCGCAGAGCTGTGATATATACAGATTTTCTTTTGTTTTTTCTTCTCCCTT	2137
DB	1800	GATACCTTAGAGGCGCAGAGCTGTGATATATACAGATTTTCTTTTGTTTTTTCTTCTCCCTT	1859
QY	2138	TTCATATGCTGTTTTCTATGTTTCCTGCTCTATATTTGATGTTTCATTTTGTAGTCTCTTG	2197
DB	1860	TTCATATGCTGTTTTCTATGTTTCCTGCTCTATATTTGATGTTTCATTTTGTAGTCTCTTG	1919
QY	2198	TAATAAGCCTGTGTTTTCTATTTGAAATTTCTGAAACGCTTCCCTTAAACGCTATTTGGCTCCC	2257
DB	1920	TAATAAGCCTGTGTTTTCTATTTGAAATTTCTGAAACGCTTCCCTTAAACGCTATTTGGCTCCC	1979

QY	2258	TTAAAACTGAACATCTCTCAATTTTGTGAT	2288
Db	1980	TTAAAACTGAACATCTCTCAATTTTGTGAT	2010
RESULT 3			
US-10-425-114-10724			
; Sequence 10724, Application US/10425114			
; Publication No. US2004003488A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yibua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(5313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 10724			
; LENGTH: 2011			
; TYPE: DNA			
; ORGANISM: Glycine max			
; FEATURE:			
; OTHER INFORMATION: Clone ID: 700963303_FLI			
US-10-425-114-10724			
Query Match 86.6%; Score 1982; DB 13; Length 2011;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1982; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	307	ATCCAAATATGACGGAGGAACGAACTGCGGAAGACACGCTGFGTGCAGCTGCTCGA	366
Db	1	ATCCAAATATGACGGAGGAACGAACTGCGGAAGACACGCTGFGTGCAGCTGCTCGA	60
QY	367	CTGCGTCAATCCCTTACATCGACGACCCCAAGACCGCGGCTTCCAGGCTGTCG	426
Db	61	CTGCGTCAATCCCTTACATCGACGACCCCAAGACCGCGGCTTCCAGGCTGTCG	120
QY	427	ACGCTGTTACGAGCTCGACTCGCTCACCCGCAAGACGCTCACCTCGGCTCTGCTACAC	486
Db	121	ACGCTGTTACGAGCTCGACTCGCTCACCCGCAAGACGCTCACCTCGGCTCTGCTACAC	180
QY	487	CACCAACCCCGGCTCGCTCCGCGCGCTTCCGCGACCTCGAGTGCCTCAAGG	546
Db	181	CACCAACCCCGGCTCGCTCCGCGCGCTTCCGCGACCTCGAGTGCCTCAAGG	240
QY	547	CAAGCCCGGAGCGGCTGATGTTCACTTGATACCGAGGATGGGGCGGACAGTCACTCC	606
Db	241	CAAGCCCGGAGCGGCTGATGTTCACTTGATACCGAGGATGGGGCGGACAGTCACTCC	300
QY	607	CTGGGTCAAAGAGATTCTCAGTACTTTCGATTGCTTCAAGAGCCTCCACTTCCGCGCAT	666
Db	301	CTGGGTCAAAGAGATTCTCAGTACTTTCGATTGCTTCAAGAGCCTCCACTTCCGCGCAT	360
QY	667	GATTGTCAAGGATTCGATTCAGAAATCTCGTCTGTCGCGGCTCAAGTCTTCAAGC	726
Db	361	GATTGTCAAGGATTCGATTCAGAAATCTCGTCTGTCGCGGCTCAAGTCTTCAAGC	420
QY	727	TCTCAAGCTTGACAGTGTCCGGTTTACCACCGATGGTCTTTTCCATATCGGTGCTT	786
Db	421	TCTCAAGCTTGACAGTGTCCGGTTTACCACCGATGGTCTTTTCCATATCGGTGCTT	480
QY	787	TTGCAAGATTAAAGAGCTCTTGTTTGGAGGAAGAGCTCAATTTCTTGAGGAAGACGAGA	846
Db	481	TTGCAAGATTAAAGAGCTCTTGTTTGGAGGAAGAGCTCAATTTCTTGAGGAAGACGAGA	540
QY	847	ATGGCTACACGAGCTTCTGTTGAATAATACAGTCTTGAGACTCTCAATTTTACTTGAC	906
Db	541	ATGGCTACACGAGCTTCTGTTGAATAATACAGTCTTGAGACTCTCAATTTTACTTGAC	600
QY	907	AGACATTGCTGTTGTGAAGATTGAGGACCTTGAACCTTTTAGCTTAAATAATGCCCACTT	966
Db	601	AGACATTGCTGTTGTGAAGATTGAGGACCTTGAACCTTTTAGCTTAAATAATGCCCACTT	660
QY	967	AGTGTCTGTGAACACTTACTGACTGTGAATCTTGTGAACCTTCTTTTAAGCATGC	1026
Db	661	AGTGTCTGTGAACACTTACTGACTGTGAATCTTGTGAACCTTCTTTTAAGCATGC	720
QY	1027	CTCTGGCTGGAAGAGTTTGTGAGGCGACCTTCAACAGAGGAAACGAGAAAGATCTCTGC	1086
Db	721	CTCTGGCTGGAAGAGTTTGTGAGGCGACCTTCAACAGAGGAAACGAGAAAGATCTCTGC	780
QY	1087	TATATCATTTACCAGCAAAAGTTATGTGATTTGTTTAAATATATTTTGGAAAGATGATTT	1146
Db	781	TATATCATTTACCAGCAAAAGTTATGTGATTTGTTTAAATATATTTTGGAAAGATGATTT	840
QY	1147	GCCCATTTGTTTCATGTTTGCAGCGTACTTAAATAATTTGATCTCTCTATGCAATGCT	1206
Db	841	GCCCATTTGTTTCATGTTTGCAGCGTACTTAAATAATTTGATCTCTCTATGCAATGCT	900
QY	1207	AGACACGAGGATCAATGTATGTTAATCCAAAGGTGTCCTGGAAGTCTTGGAGC	1266
Db	901	AGACACGAGGATCAATGTATGTTAATCCAAAGGTGTCCTGGAAGTCTTGGAGC	960
QY	1267	AAGGAATGTAATTGGAGATAGAGGTTAGAGGTTCTTGGTCTTGTAAAGAGCTTAA	1326
Db	961	AAGGAATGTAATTGGAGATAGAGGTTAGAGGTTCTTGGTCTTGTAAAGAGCTTAA	1020
QY	1327	AAGGCTTAGGATTGAAAGGGCGATGATGATCAAGGAATGGAGGATGAAGAGTACTGT	1386
Db	1021	AAGGCTTAGGATTGAAAGGGCGATGATGATCAAGGAATGGAGGATGAAGAGTACTGT	1080
QY	1387	GTCCCATAGAGGGCTAATAGCTTGTCAAGGCTGTTTCAAGGCTTGAATCATGGCTGT	1446
Db	1081	GTCCCATAGAGGGCTAATAGCTTGTCAAGGCTGTTTCAAGGCTTGAATCATGGCTGT	1140
QY	1447	TTATGTGTCTGATATTCAATGATCTCTGGAACATATTGGAACCTCACTTGAAGAACT	1506
Db	1141	TTATGTGTCTGATATTCAATGATCTCTGGAACATATTGGAACCTCACTTGAAGAACT	1200
QY	1507	CTGTGATTTTCGCTTGTGTTGCTTACCATTAAGAGAGATACTGATTTGCCACTTGA	1566
Db	1201	CTGTGATTTTCGCTTGTGTTGCTTACCATTAAGAGAGATACTGATTTGCCACTTGA	1260
QY	1567	CAATGGGCTGAGGGCTCTACTGAGGGGCTGTGACAAAGCTGAGGAGTTTGTCTATATCT	1626
Db	1261	CAATGGGCTGAGGGCTCTACTGAGGGGCTGTGACAAAGCTGAGGAGTTTGTCTATATCT	1320
QY	1627	CAGGCTGGCGGTTGACTGATGAGGCTTGGTTTACATTTGGACAATAACAGTCCAAATGT	1686
Db	1321	CAGGCTGGCGGTTGACTGATGAGGCTTGGTTTACATTTGGACAATAACAGTCCAAATGT	1380
QY	1687	GAGATGGATGCTGCTGTTTATGTTGGGGAGTCTGATGAGGGCTTGGAGTTTGGCTTAA	1746
Db	1381	GAGATGGATGCTGCTGTTTATGTTGGGGAGTCTGATGAGGGCTTGGAGTTTGGCTTAA	1440
QY	1747	GGGGTGTCTAGCTCTTCAAGAACTTGAAGAGGCTTATTTTTCAGTGAAGCTGC	1806
Db	1441	GGGGTGTCTAGCTCTTCAAGAACTTGAAGAGGCTTATTTTTCAGTGAAGCTGC	1500
QY	1807	ACTTGTGTTGGCTGCAACACAAATTCAGTCTTCTTAGGTACTTGTGGGTGCAAGGTTATGG	1866
Db	1501	ACTTGTGTTGGCTGCAACACAAATTCAGTCTTCTTAGGTACTTGTGGGTGCAAGGTTATGG	1560
QY	1867	TGTATCTCATCTGAGAGTGTATCTTTTGGTAATGGCTCGACCTTTTGGAAATTTGAGTT	1926
Db	1561	TGTATCTCATCTGAGAGTGTATCTTTTGGTAATGGCTCGACCTTTTGGAAATTTGAGTT	1620
QY	1927	GATTCCTCTAGAAAGGCTGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTGAGCA	1986
Db	1621	GATTCCTCTAGAAAGGCTGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTGAGCA	1680



1987 TCCTGCTCATATCTTGCATATATCTCTTGCAGGCGAGATCAGATTTTCCAGATAC 2046  
1681 TCCTGCTCATATCTTGCATATATCTCTTGCAGGCGAGATCAGATTTTCCAGATAC 1740  
2047 TGTGTGCTTTTGGACACTGCCACATCGTGTGATACCTAGAGGCGAGCTGTGTATATA 2106  
1741 TGTGTGCTTTTGGACACTGCCACATCGTGTGATACCTAGAGGCGAGCTGTGTATATA 1800  
2107 TACAGATTTCTTTTGTGTTTCTTCTCCCTTTCAATATGCTGTTTCTATGTTCTGCTCT 2166  
1801 TACAGATTTCTTTTGTGTTTCTTCTCCCTTTCAATATGCTGTTTCTATGTTCTGCTCT 1860  
2167 ATTTGTAGTTCATTTTAGACAAATAGTCTTGTATATAGCCCTGTTTTCATTTGAAATTC 2226  
1861 ATTTGTAGTTCATTTTAGACAAATAGTCTTGTATATAGCCCTGTTTTCATTTGAAATTC 1920  
2227 TGAACGCTTCCCTTAAAGCTTATTTGGCTCCCTTAAACAACTGAACATTTCTCAATTTTGTGA 2286  
1921 TGAACGCTTCCCTTAAAGCTTATTTGGCTCCCTTAAACAACTGAACATTTCTCAATTTTGTGA 1980  
2287 AT 2288  
1981 AT 1982

RESULT 4  
US-10-310-154-152  
; Sequence 152, Application US/10310154  
; Publication No. US20030233670A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Chomet, Paul S.  
; APPLICANT: Adams, Thomas H  
; APPLICANT: Ruff, Thomas G.  
; APPLICANT: Agarwal, Ameeta K.  
; APPLICANT: Ahrens, Jeffrey E.  
; APPLICANT: Ball, James A.  
; APPLICANT: Banu, G.  
; APPLICANT: Bell, Erin  
; APPLICANT: Boddupalli, Raghava  
; APPLICANT: Deikman, Jill  
; APPLICANT: Deng, Molian  
; APPLICANT: Dong, Jinchuo  
; APPLICANT: Duff, Stephen M.  
; APPLICANT: Galligan, Meghan M.  
; APPLICANT: Hinchey, Brenda S.  
; APPLICANT: Huang, Shihshieh  
; APPLICANT: Johnson, G. Richard  
; APPLICANT: Jung, Vincent  
; APPLICANT: Kretzmer, Keith A.  
; APPLICANT: Laccetti, Lucille B.  
; APPLICANT: Lai, Chao-Qiang  
; APPLICANT: Lee, Gary  
; APPLICANT: Lin, Jie-Yi  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lu, Bin  
; APPLICANT: Luethy, Michael M.  
; APPLICANT: Lund, Adrian  
; APPLICANT: Madson, Linda L.  
; APPLICANT: Malloy, Kathleen A.  
; APPLICANT: McKiel, Christine L.  
; APPLICANT: Miller, Philip W.  
; APPLICANT: Padmavathi, Manchikanti  
; APPLICANT: Parnell, Laurence D.  
; APPLICANT: Start, William G.  
; APPLICANT: Tennesen, Dan  
; APPLICANT: Vidya, K.R.  
; APPLICANT: Wang, Haiyun  
; APPLICANT: Xin, Zhanguo  
; APPLICANT: Xu, Nanfei  
; APPLICANT: Yang, Chunzhi  
; APPLICANT: Zeng, Xiaoping  
; APPLICANT: Zhang, Qiang

APPLICANT: Zhao, Yajuan  
APPLICANT: Zhou, Li  
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants  
FILE REFERENCE: 38-15(5796)B  
CURRENT APPLICATION NUMBER: US/10/310,154  
CURRENT FILING DATE: 2002-12-04  
PRIOR APPLICATION NUMBER: 60/337,358  
PRIOR FILING DATE: 2001-12-04  
NUMBER OF SEQ ID NOS: 736  
SEQ ID NO 152  
LENGTH: 2178  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (130)..(1923)  
OTHER INFORMATION:  
US-10-310-154-152

Query Match 26.7%; Score 609.8; DB 16; Length 2178;  
Best Local Similarity 61.9%; Pred. No. 1.5e-176;  
Matches 1076; Conservative 0; Mismatches 642; Indels 21; Gaps 6;

Qy 345 GTGTGCTGAGCTGCTCTGAGCTGCTCATCCCTTACATCGAGCCCAAGGACCGCG 404  
Db 191 GCGTCCCGGAGGCGCTGCACCTCGTGTGGCTACATGGACGCCCGGAGACCGGG 250  
Qy 405 ACGCGGTTTCCAGGCTGTGCGAGCTGTGACGAGCTCGACTCGCTCAACCGCAAGCAG 464  
Db 251 AGGCGGCTGCTGCTGTGCGGCTCTGGCACCGCATCGAGCGCTCTCGGCGAAGCAG 310  
Qy 465 TCACCATGCGGCTCTGTCTACACACACCCCGGCTGCGCTCCGCGGCTTCCGCGACC 524  
Db 311 TCACCGTGGGCTTCTGTACGCGCGCGCGGCTGCTCGCCAGGTTCCGCGGCG 370  
Qy 525 TCGAGTGGCTCAAGCTCAAGGGCAAGCCCGAGCGCAATGTTCAACTTGTATACCCGAGG 584  
Db 371 TCGAGTGGCTCGGCTCAAGGGGAGGCCACCGCGCGCATGTACGGGCTCATCCCGAAG 430  
Qy 585 ATTGGGCGGACAGCTCACTCCCTGGGTCAAAGAGATTTCTCAGTACTTCGATTCGCTCA 644  
Db 431 ACTTTGGGCGCTACGGGCGGCTGCTGGTCCGCGAGCTCGCGCGCCCTTGACTGCTCA 490  
Qy 645 AGAGCTCCACTTCGCGCGCATGTGTCAAGGATTCGATCTTCAGATCTTCGCTCGTG 704  
Db 491 AGGCGCTCCACTCGCTGCGCATGACCGTCACCGAGGAGACATCGCGCTGCTCCAGC 550  
Qy 705 ACCGCGGTGAGTCTTCAAGCTCTCAAGCTTGAAGTGTCCGCTTTCACACCGGATG 764  
Db 551 CGCGCGGCCACATGCTACAGGCGCTCAAGCTCGACAGTGTCTCCGCTTCTCAACAGAG 610  
Qy 765 GTCTTTTCCATATCGGTCGCTTTTTCAGAGATTTAAGAGTCTTGTGTTTGGAGGAAAGT 824  
Db 611 CCTTCGCGCTCGTCCGCGGCTCTCTGAGAACCTTGTTCCTGGAAGATGTA 670  
Qy 825 CAATTTTGAAGAGACGAGAAATGCGCTACAGACTTGTCTTGTGAATATACAGTTCTTG 884  
Db 671 TAATTTGAGATGAGAGGAGTGAATGGCTCCATGAATCTGCTGTCAACAATCTGTTCTGG 730  
Qy 885 AGACTCTCAATTTTACTTGAAGAGATGCGCTACAGACTTGTCTTGTGAATATACAGTTCTTG 944  
Db 731 TGACACTGAACCTTTTACATGACAGAACT---CAAGTGGAGCGCTGCTGATCTGGAGCTTC 787  
Qy 945 TAGCTAAAAATTCGCCCAACTTAGTGTCTGTGAATCTTACTGACTGTGAATCTGGATC 1004  
Db 788 TTGCAAGAGACTGTAAATCATTTCTTTGAAGATGGTGAATGCTGTGATCTTTTCAGATC 847  
Qy 1005 TTGTGAACCTTTTAAAGCATGCTTGGCTGGAAGAGTTTGTGGAGGACCTTCAACG 1064  
Db 848 TGATTGGGTTTTTCCAAACCTCCAAAGCATTCGAAGAAATTCGCGGTGTGCTTTTGG 907  
Qy 1065 AGGAACACAGA-----AAGATCTCTGCTATATCATTTACAGCAAGATTTATGTCGAT-- 1116

Tue Apr 20 10:32:12 2004

us-10-009-791-21.rnpg

Db 908 AAGTTGAGAGTACACCAAGTACGAAAGGTCAATTTCCACCTAGGCTATGCTTCTTAG 967  
Qy 1117 -GGGTTTAAACATATATTGGAAGATCAGTTGGCCATTGTTTCATGTTTGCAGCGGTAC 1175  
Db 968 GGGGCTTACCTTCATGGTAAACAGATGCCAGTTATCTTCCGATTTCCAAATGC 1027  
Qy 1176 TAAAAAATGGATCTCCTCTATGCAATGCTAGACAGGAGATCATTTGATGTTATATCC 1235  
Db 1028 TTAAGAACTGGACTTGCAGTTCACTTTCTCCACCTGAAAGACCACTGTGAGCTTATTG 1087  
Qy 1236 AAAGGTGTCAAAATCTGGAAGTCTTGGAGACAGGAATGTAATTTGGAGATAGAGGTTAG 1295  
Db 1088 CTAATGCCCCAACCTAAGTCTTCGAGGTGAGGAATGTGATTTGGGATAGAGACTAG 1147  
Qy 1296 AGGTTCTTGGTGGTGTGTAAGAGGCTTAAAGGCTTAGGATTTGAAAGGGGCGATGATG 1355  
Db 1148 AAGTTGTTGCTGTACATGCAAGACTACGAGGCTCAGAGGCTGAGGCTGGCGACGATG 1207  
Qy 1356 ATCAAGGAATGGAGATGAAGAGTACTGTGTCCATAGAGGGCTAATAGCCTTGTAC 1415  
Db 1208 ATCTGGCCAAAGAGCAGGAGGAGG---AGTCTCTCAGATAGTTTGACAGCCGTAGCCG 1264  
Qy 1416 AGGGCTTTCAGAGCTTGAATACATGCTGTTTATGCTGTGATATTACAAATGCATCTC 1475  
Db 1265 TTGGCTGCCGGAACAGTACATAGCTGCTATGCTATCTGATATCAAAATGGGGCGC 1324  
Qy 1476 TGAACATATTGGAACCTCACTTGAAGAACCTCTGTGATTTTGGCCTTGTGTTGAC 1535  
Db 1325 TGAATCTCATTTGGCACTTCTGCAAGATCTCTATGATTTCCGGCTTGTCTACTTGACA 1384  
Qy 1536 ATGAAGAGAAGATACTGATTTGCCACTGTGAACAATGGGGGTAGGGCTCTACTGAGGGCT 1595  
Db 1385 AACAGAAGAGATAACAGACCTGCCACTTGACACAGGGCTCGCTGCCCTACTGAGGAAT 1444  
Qy 1596 GTGCAAGCTGAGGATTTGCTCTATATCTCAGGGTGGCGGCTTGAAGTATGAGGC 1655  
Db 1445 GCGTCAAACTTCGAGGTTTGGCTTCTACTGAGACCGGAGGGCTTTCAGACGTAGGCC 1504  
Qy 1656 TTGGTTACATTTGGAACAATACAGTCCAAATGTGAGATGATGCTGTGTTGTTATGTTGGG 1715  
Db 1505 TCGGTTACATTTGGAATACTACAGTGGGAACATCCAATACATGCTGCTCGGCAAGCTTGGTG 1564  
Qy 1716 AGTCTGATGACGGCTTTTGGAGTTCGCTTAAGGGGTGCTAGTCTTCAGAACTTGAA 1775  
Db 1565 AATCTGACAAAGGATTAATGATCCAGTTTGAACGGGATGCAACCTGCGGAAGCTCGAGC 1624  
Qy 1776 TGAGAGGTGTTTATTTTTCAGTGAAGTGCATCTGCTGTTGGCTGCAACACAAATGACCTT 1835  
Db 1625 TGAGGGGCTG---CTGCTTACGAGGAGAGCTCTGGCAGTGGCCGTCTCCAGATGCCCT 1681  
Qy 1836 CTCTTAGGTACTTGTGGGTGCAAGTTATGTTGTATCTCCATCTGACGCTGATCTTTTGG 1895  
Db 1682 CGCTGAGGTACATATGTTGTCAGGGTACAGAGCTCTCGAAGGGCCAGGACCTCATGC 1741  
Qy 1896 TAATGGCTGACCTTTTGGAAATTTGAGTT---GATTCCTTTAGAAAGTGGCTACGA 1952  
Db 1742 TCATGGCCAGCGGTACTGGAACATTTGAATTCGGCGCTCCCATTCGAGAGGCGCTTATC 1801  
Qy 1953 ATACCAATCCAGATGAGACTGTAGTTGTTGAGCATCTGCTCATATTTCTTGSCATATTAT 2012  
Db 1802 GGGTGAATGGCAGATGGAACGCTTGTGTTGACACCCATGCCAGGCTCTTGGCTATTACT 1861  
Qy 2013 CTCTTGAGGGCAGAGATCAGATTTTCCAGATATGTTGTCCTTTGGACACTGCCACA 2071  
Db 1862 CCCTTGTGGAAGAGGCGGCGGACTGCCCTCAGTGGCTGGTGAATTTGACCCCTCGCTCA 1920

RESULT 5  
US-10-310-154-151  
; Sequence 151, Application US/10310154  
; Publication No. US20030233670A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D

; APPLICANT: Chomet, Paul S.  
; APPLICANT: Adams, Thomas H  
; APPLICANT: Ruff, Thomas G.  
; APPLICANT: Agarwal, Ameeta K.  
; APPLICANT: Ahrens, Jeffrey E.  
; APPLICANT: Ball, James A.  
; APPLICANT: Banu, G.  
; APPLICANT: Bell, Erin  
; APPLICANT: Boddupalli, Raghava  
; APPLICANT: Deikman, Jill  
; APPLICANT: Deng, Molian  
; APPLICANT: Dong, Jinchuo  
; APPLICANT: Duff, Stephen M.  
; APPLICANT: Galligan, Meghan M.  
; APPLICANT: Hinchey, Brenda S.  
; APPLICANT: Huang, Shihshieh  
; APPLICANT: Johnson, G. Richard  
; APPLICANT: Jung, Vincent  
; APPLICANT: Kretzmer, Keith A  
; APPLICANT: Laccetti, Lucille B.  
; APPLICANT: Lai, Chao-Qiang  
; APPLICANT: Lee, Gary  
; APPLICANT: Lin, Jie-Yi  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Lu, Bin  
; APPLICANT: Luethy, Michael M.  
; APPLICANT: Lund, Adrian  
; APPLICANT: Madson, Linda L.  
; APPLICANT: Malloy, Kathleen A.  
; APPLICANT: McKiel, Christine L.  
; APPLICANT: Miller, Philip W.  
; APPLICANT: Padmavathi, Manchikanti  
; APPLICANT: Parnell, Laurence D.  
; APPLICANT: Start, William G.  
; APPLICANT: Tennesen, Dan  
; APPLICANT: Vidya, K.R.  
; APPLICANT: Wang, Haiyun  
; APPLICANT: Xin, Zhanquo  
; APPLICANT: Xu, Nanfei  
; APPLICANT: Yang, Chunzhi  
; APPLICANT: Zeng, Xiaoping  
; APPLICANT: Zhang, Qiang  
; APPLICANT: Zhao, Yajuan  
; APPLICANT: Zhou, Li  
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants  
; FILE REFERENCE: 38-15(52796)B  
; CURRENT APPLICATION NUMBER: US/10/310,154  
; CURRENT FILING DATE: 2002-12-04  
; PRIOR APPLICATION NUMBER: 60/337,358  
; PRIOR FILING DATE: 2001-12-04  
; NUMBER OF SEQ ID NOS: 736  
; SEQ ID NO 151  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (139)..(1911)  
; OTHER INFORMATION:  
US-10-310-154-151

Query Match 24.5%; Score 561.2; DB 16; Length 2419;  
Best Local Similarity 61.3%; Pred. No. 1.7e-161;  
Matches 1000; Conservative 0; Mismatches 603; Indels 27; Gaps 5;  
Qy 327 GGAACTGGGAAGACACAGTGTGTCGACTGCTCGACTGCGTTCATCCCTTACATCG 386  
Db 176 GGCGGGTGTGAGCTTTGGGATCCCGGACACCGCGCTAGTGTATGGGGTACGTGG 235  
Qy 387 ACCACCCCAAGGACCGCGCGCTTTCCAGGTGTGTCGACGCTGTGTACGACTCGACT 446  
Db 236 AGGATCCCTGGGACCGCGCGCATCTCGTGTGTGTGCGGACTGTGTGCGGCGTGCAGC 295



981 TTACTGACTGTGAATACCTGATCTTGTGAACCTTTCTTTAAGCATGCTCTGCGCTGGAAG 1040  
 Db TGGGTGACTGTGATCTTTGAGACTGATTTGGTCTTTTCCAACTCCAAAGCATTTGCAAG 237  
 1041 AGTTTTGTGAGGACCTTACACAGAGAACAGAG-----AGATATCTCTGTATATCAT 1094  
 Db AATTGCGCGGTGGTGGCTTTTGAAGTGTGAGAGTACACCAAGTACGAAAGGTCAATTT 297  
 1095 TACAGCAAGTATATGTCGATT---GGTTTAAATATATTTGGAAGAAATGAGTTGCCCA 1151  
 Db TCCACACCTAGGCTATGCTTTAGGGGTCTTACCTTCATGGGTGTAAGAGAGATGCCAG 357  
 1152 TTGTGTTCACTTTGCGAGCCGTACTAAAAAATTTGGATCTCTCTATGCAATGCTAGACA 1211  
 Db TTATCTTTCCGTATTTCCACATGCTTAAAGAACTGGACTTCAGTTCACTTTCTCTACCA 417  
 1212 CGGAGGATCATTTGATTAATCAAGAGTGTCCAAATCTGGAAGTCTTCCAGACAGGA 1271  
 Db CTGAAGACCACTGTCAAGCTTATGCTAAATGCCCAACCTAAGTGTCTCGAGGTGAGGA 477  
 1272 ATGTAATTTGGAGATAGAGGCTTAGAGGTTCTTTGGTCTGTTGTTAAGAGGCTAAAGGC 1331  
 Db ATGTGATTTGGGATAGAGGCTAGAGGTTGTTGCTGTACATGCAAGAGCTACGAAGGC 537  
 1332 TTAGGATTTGAAGGGCGCATGATCAAGAAATGAGAGATGAAGAGTACTGTCTCC 1391  
 Db TCAGAATTTGAGCGTGGCAGCATGATCTCTGGCCAAAGAGAGCAGGAGG---AGTCTCTC 594  
 1392 ATAGAGGCTATAGCTTGTTCACAGGCTGTTTCAGAGCTTCAATACATGCTGTTTATG 1451  
 Db AGATAGGTTTACAGCGCGTAGCGGTTGGCTGCGGCACTAGAGTACATAGTGTCTGCTATG 654  
 1452 TCTCTGATATTAATAATGATCTCTGGAAATATTTGGAACTCACTTTGGAAGAACCTCTGTG 1511  
 Db TATCTGATATCAAAATGGGCGCTGGATCCATTTGGCACTTCTTCTGCAAGATCTCTATG 714  
 1512 ATTTGCGCTTGTGTTGCTTGACCATGAGAGAGATCACTGATTTGGCCACTTGCACATG 1571  
 Db ATTTCCGGCTTGTCTTACTTGAACACAGAGAGATCACTGACCTTGCACCTTGCACAG 774  
 1572 GGGTGAAGGCTTACTGAGGCGGTGTGCAAGCTGAGAGATTTGCTCTATATCTCAGCG 1631  
 Db GCGTCCGTGCTTACTGAGGACTGCGGTCAACTTGGAGGTTGCTTCTTCTTCTTCTGAGAC 834  
 1632 GTGGCGGCTTACTGATGTAGGCTTTGTTTCAATGCAATACATGCTCAATGTGAGAT 1691  
 Db CGGAGGGCTTTTCAAGCTAGGCTCGGTTTACATTTGACATATACAGTGGGAAATCCCAAT 894  
 1692 GGATGCTGCTGTTTATGTTGGGAGTCTGATGCGAGGCTTTTGGAGTTGCTTAAGGGGT 1751  
 Db ACATGCTGCTCGGCAAGCTTGTGTAATCTGCAACCGGATTTGATCCAGTTTGGCAACGGAT 954  
 1752 GTCTAGTCTTCAAGAACTTTGAATAGAGGCTGTTTATTTTTCAGTGAACGTGCACTTG 1811  
 Db GCACCAACCTCGGAGAGCTGAGCTGAGGGGCTG---CTGCTTCCAGGAGCGAGCTCTGG 1011  
 1812 CTGTGGCTGCAACACAAATTTGACTTCTCTTAGGTAATTTGTTGGGTGCAAGGTTATGTTAT 1871  
 Db CAGTGGCGTGTCTCCAGATGCTCTGCTGAGTACATATGGGTGCGAGGGGTACAGAGCCT 1071  
 1872 CTCATCTGAGAGCTGATTTTGTGTAATGCTCGACCCCTTTTGGAACTGAGTT---GA 1928  
 Db CTCGAACCGGCGCAGGACCTTCTGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1131  
 1929 TTCTTCTTGAAGAGGTGCTACGAATACCAATCCAGATGAGACTGTAGTTTGTTCAGCATC 1988  
 Db TCCCATCTCCGAGAGCGCTTATCGGCTGATGGCAGATGGACAGCTTGTGTTGACACCC 1191  
 1989 CTGCTCATATTTCTTCATATTTCTCTTTGAGGCGAGAGATCAGATTTTCCAGTACTG 2048  
 Db ATGCCAGGCTCTTGGCTATTACTCTCTTGTGGAAGGAGGCGGAGCTGCGCTCAGTGGC 1251  
 2049 TTGTGCTTTGGACACTGCCACA 2071

1252 TGGTGACTTTGACCCCTGGCTCA 1274  
 Db  
 RESULT 7  
 US-10-425-114-8867  
 ; Sequence 8867, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 8867  
 ; LENGTH: 1325  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700800510\_FLI  
 US-10-425-114-8867  
 Query Match 17.6%; Score 403.8; DB 13; Length 1325;  
 Best Local Similarity 60.2%; Pred. No. 4.6e-113;  
 Matches 726; Conservative 0; Mismatches 467; Indels 12; Gaps 3;  
 QY 923 AAGATTGAGGACCTTGAACCTTTTAGCTAAAAAATTCGCCCAACTTAGTGTGTGAAACTT 982  
 Db 32 AGGACAGTGTGATTAGAGCTTCTTGCAGAGAGCTGCAAGTCACTGATTTCTTGAAGATT 91  
 QY 983 ACTGACTGTGAATATCTGATCTTGTGAACTTCTTTAAGCATGCTCTGCGCTGGAAGAG 1042  
 Db 92 GCGGATTTGTGATCTTTTCAGATTTGATAGGTTCTTCCAACTCTGCCACATCATCTGGAAGA 151  
 QY 1043 TTTTGTGAGGACCTTACACAGGAGAACAGAA-----AGATATCTCTCTATATCATTTA 1096  
 Db 152 TTTTGTGAGGAGACATTTAATGGCAGAGGGAACCTCACCAAGTATGGGATGTAAATTT 211  
 QY 1097 CAGCAAAAGTATGCTGATTTGAGTTTAAATATATTTGGAAGAAATGAGTTGCCATTGTG 1156  
 Db 212 CCATCAAGAAATATGCTCTTGGGACTTACTTTTCATGGGTGCAAAATGAAATGCTATATA 271  
 QY 1157 TTTTGTGAGGCGGTACTTAAAGAAATTTGGAATCTCTCTATGCAATGCTAGACACGGAG 1216  
 Db 272 TTTTCTTTTCTGCGATCTTAAAGAGAGCTGGATTTTGCAGTACACTTTCTCTCACCGAA 331  
 QY 1217 GATCATTTGATGTTAATCCAAAGGCTGTCAAATCTGGAAGTCTCTGAGACAGGAATGTA 1276  
 Db 332 GACCAATGCGAGCTCAATGCAAAATGCCGAATCTTACTAGTTCTCGGGTGAGGAATGTG 391  
 QY 1277 ATTGGAGATAGAGGTTTAGAGGTTCTTGGTCTGTTGTTAAGAGGCTTAAAGAGCTTAGG 1336  
 Db 392 ATTGGGATAGAGGATTTAGAGTTTGTTCAGATACGTCGCAAGAGGCTTCCAAAGGCTCAGA 451  
 QY 1337 ATTGAAGGGGCGATGATGATCAAGAAATGAGAGGATGMAAGGTAAGTCTGTGCCATAGA 1396  
 Db 452 ATTGAGCGAGGAGCAGATGAAGAGGTTGTCAGAGGAGGAGCGAGGGGTCTCTCAAGTG 511  
 QY 1397 GGGCTAATACCTTTGTTCACAGGCTGTTTTCAGAGCTTGAATATACATGCGTGTATGTGCT 1456  
 Db 512 GGGTGTGACGCTATAGCGGTAGGCTGTCTGTAAGTGAATATACATAGCTGCTATGTGCT 571  
 QY 1457 GATATTAACAATGATCTCTGGAACATATTTGGAATCACTTGAAGAACCTCTGTGATTT 1516  
 Db 572 GATATTAACCAATGGGCTTTGGAATCTATCGGACATTTCTGCAAAAAAATCTATGACTTC 631



Tue Apr 20 10:32:12 2004

us-10-009-791-21.rnpb

```

1919 ATTGAGTTGATTCCTTCTAG---AAAGGTGGCTACGAATACCAATCCAGATGAGACTGTA 1975
1973 ATTGAATTCGGCCCTCCAGTCCCGAGAGCGCTTATCGGGTATGATGGCAGATGGACAGCGCT 1132
1976 GTTGTGTAGACATCTGCTCATATTTCTTGCATATTTATCTTTCAGAGGCGAGATCAGAT 2035
1133 TGTGTGTACACTCATGCCAGGTCCTTTCGCTATTTACTCCCTTGTCTGGAAGGAGCGCGGAC 1192
2036 TTTCCAGATCTGTTGTGCTTGGACTTGGCACTGCCACA 2071
1193 TGGCCCTCAGTGGCTGTGACTTTCGACCTTCGCTCA 1228

RESULT 9
US-10-424-599-125710
; Sequence 125710, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 125710
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(669)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84525C.1
US-10-424-599-125710

Query Match 14.1%; Score 323; DB 13; Length 669;
Best Local Similarity 81.3%; Pred. No. 2.5e-88;
Matches 374; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1628 AGCGTGGCGGCTTGACTGATGTAGGCTTGGTATGATGAGAGAGAGTCCCAATGTG 1687
DB 1 AGACCTGGGGGATGACTGATGTGGCTAGGTACGTAGGCAATACAGCCCAACGTG 60
QY 1688 AGATGATGCTGCTTGGTTATGTGGGGAGTCTGATGAGGCTTTTGGAGTTGCTAAG 1747
DB 61 AGATGATGCTTCTTGGTTATGTGGGAGAGACTGATGAGGCTTTTGGAAATCTCTAAG 120
QY 1748 GGGTGTCTAGTCTTCAGAACTTGAATGAGAGGGTGTATTTTTCAGTGAACGTGCA 1807
DB 121 GGTGCCCCAGTCTTCAGAACTTGAATGAGAGGGTGTCTCTTTCAGTGAATGCA 180
QY 1808 CTTGCTGTGCTGCAACACAAATGACTTCTCTTAGTGA CTTGTGGGTGCAAGTTAGT 1867
DB 181 CTAGCTATTCGCACTCACTGAAATCTCTCAGGTACCTATGAGGTGCAAGGCTATAGT 240
QY 1868 GTATCTCCATCTGCACTGATCTTTTGGTAAATGCTGACCTTTTGGCACTTGAAGTTG 1927
DB 241 GCATCTGATCTGACGCGATCTTCTGCAATGCTTCTGCAATGCTTCTGCAATGCTTCTG 300
QY 1928 ATTCTCTCTAGAGGCTGCTGCAATACCAATCCAGATGAGACTGTAGTTGTGAGCAT 1987
DB 301 ATTCTCTCTAGAGGCTGCTGCAATACCAATCCAGATGAGACTGTAGTTGTGAGCAT 360
QY 1988 CCGTCTCATATCTTGCATATATTTCTCTCCAGGCGAGATCAGATTTTCCAGATCT 2047
DB 361 CTGCTCATATCTTGCATATATTTCTCTCTGCGGCCAGAAACAGATTTTCCAGATCC 420
QY 2048 GTTGTGCTTTGGACACTGCCACATGCGTTGATACCTAGA 2087

```

```

DB 421 GTAATACCTTTGGATCCTGGAAACATATATGTTGCACACCTTAAA 460

RESULT 10
US-10-424-599-125711
; Sequence 125711, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 125711
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(744)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84526C.1
US-10-424-599-125711

Query Match 13.6%; Score 311.6; DB 13; Length 744;
Best Local Similarity 75.1%; Pred. No. 9.1e-85;
Matches 431; Conservative 0; Mismatches 134; Indels 9; Gaps 3;

QY 1637 GGGTTGACTGATGTAGGCTTGGTTACAT-TGGACAATACAG-TCCAAATGTGAGATGGA 1694
DB 2 GGATTGACTGATGTAGGCTCGGTTACATAGGGGCAATACAGCCCCAACGCTGAGATGGA 61
QY 1695 TGCTGCTTGGTTATGTGGGGGAGTCTGATGAGGCTTTTGAGTTGCTTAAGGGTCTC 1754
DB 62 TGCTTCTTGGTTATGTGGGAGAGACTGATGAGGCTTTTGAAATTTCTTAAGGGTCTC 121
QY 1755 CTAGCTTTCAGAACTTGAATGAGAGGCTGTTATTTTTCAGTGAACGTGACCTTGTG 1814
DB 122 CCAGTCTTCAGAACTTTCAGATGAGGGGATGTTTCTTTTCAGTGAATGCTGACCTAGTA 181
QY 1815 TGCTGCAACAAATGACTTCTCTTAGTACTTCTGCTGGTGCAGGTTATGCTGATCTC 1874
DB 182 TTGCTGCAACTCACTGAAATCTCTCAGGTACCTATGGTGCAGGGTATGGAGCATCTC 241
QY 1875 CATCTGGAGCTGATCTTTTGGTAAATGGCTCGACCTTTTGGACATTTGATTCCTT 1934
DB 242 CATCTGGAGCTGATCTTTTGGCAATGGCTCGCCCTTATTTGGAACTTGAATTCCTT 301
QY 1935 CTAGAAAGGTGGCTAGCAATCCATCCAGATGAGACTGTAGTTGTGAGCATCTGCTC 1994
DB 302 CTAGAGCGCTTGTGTGTAGCAATCAGCAAGAGAGAGCCGCTAGTTGTGAGCACCCGCTC 361
QY 1995 ATATCTCTGCATATATTTCTTTCAGAGGCGAGAGATCAGATTTCCAGATCTGTTGTC 2054
DB 362 ATATCTTGTGATATTTACTCTTCTGCTGGCCCAAGACAGATTTTCCAGATCTGTTATAC 421
QY 2055 CTTTGGACACTGCAATGCTTGTGATACCTAGA-----GGCCAGGCTGTGATATAT 2107
DB 422 CTTTGGATCTCTGCAATATGTTGACACCTTAACTTGTACATACCACTTTTCTTCTTC 481
QY 2108 ACCAGTTTCTTTTGTGTTTCTTCTCCCTTTCATATGCTGTTTCTATGTTCTGCTCTA 2167
DB 482 TTCTTCTTCCAAAGGTTTTTTTTTCTCCCTTCAATCTTGTGTTGTAAGTGGTGGAGTG 541
QY 2168 TTTGTAGTTCATTTTAGCAATTAGTCTTGTAAAT 2201

```



Db 542 TTTTAGTTAGAGGTTTCATTTTCTCTGCAAT 575

## RESULT 11

US-10-424-599-317  
; Sequence 317, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 317  
; LENGTH: 829  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_100286C.1  
US-10-424-599-317

Query Match 12.7%; Score 290.4; DB 13; Length 829;  
Best Local Similarity 74.4%; Pred. No. 3.5e-78;  
Matches 366; Conservative 0; Mismatches 126; Indels 0; Gaps 0;  
  
QY 319 GGAGGAACGGAACGTGGGGAAGACACGTGTGGTCGACGTGTCTCGACTCGCTCATCC 378  
Db 245 GGACGGAGCGGATGACGGGAGGAGTTATCGGATGTGGTCTGACTTGGTATGCC 304  
QY 379 TTATCATGACGACCCCAAGAACCGGACCGCGTTCCTCCAGGTGTGTGACGCTGGTACGA 438  
Db 305 ATACATACACGACTCCAAAGGACCGGACCGCGTGTGCGAGGTGTGCGAGCGGTGGTACGA 364  
QY 439 GCTCGACTGCTCACCGGCAAGACGTCACATCGGCTCTGTACACCAACACCGCGG 498  
Db 365 GCTCGACTGCTCACTCGTAAACACGTCACATTTGGGCTGTGTATACCAACACCGCGG 424  
QY 499 TCGCTCCGCGCGCGCTTCGCGACCTCGAGTCGCTCAAGCTCAAGGGCAAGCCGCGAGC 558  
Db 425 CCGGCTTAAGCGCGGGTTCGCGACCTCGAGTGTGCTGAAGCTGAGGGAAGCGCGGCG 484  
QY 559 CGCAATGTTCAAATTGATACCGGAGATGGGGCGGACACGTCACTCCCTGGGTCAAAGA 618  
Db 485 GCGGATGTTCAAACCTGATACCTGAGGATGGGGAGGTTTGTGTCAAGCTTGGGTCAAGT 544  
QY 619 GATTTCTCAGTACTTCGATGCTCAGAGCTCCACTTCGCGGCGATGATTTGTCAGGA 678  
Db 545 GATCTCTCAATCTTTGATTTGTTGAAGTCGCTTCATTTTCGCGCGCATGATTTGTGAGGA 604  
QY 679 TTCGATCTTCAGAACTCTCGCTGTGACCGCGGTTCAGTCTTCAAGCTTCAAGCTTGA 738  
Db 605 TTGGAATCTTCAAGTCTAGCTGTTCGCGCGGTTCAGTACTTCAAGCACTCAGGCTGGA 664  
QY 739 CAAGTGTCCGGTTTCAACCGAGATGCTTTTCCATATCGGTGCTGCTTTTCAAGAGTTT 798  
Db 665 AAATTTGCTCTGGATTTCTCCACTGATGCGCTCTACTATAGTCAATTTAAGTGTGA 724  
QY 799 AAGAGTCTTCTT 810  
Db 725 GGCTCACTTGT 736

## RESULT 12

US-10-424-599-44612  
; Sequence 44612, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 44612  
; LENGTH: 428  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_140283C.1  
US-10-424-599-44612

Query Match 12.2%; Score 279.8; DB 13; Length 428;  
Best Local Similarity 80.9%; Pred. No. 4e-75;  
Matches 326; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
  
QY 577 ACCGAGGATGGGCGGACACGTCACTCCCTGGGTCAAAGAGATTTCTCAGTACTCGA 636  
Db 25 ACCGAGGATGGGAGGTTTGTCTACGCGGGGGTCAAGAGATCTCTCAGTACTCGA 84  
QY 537 TTGCCTCAAGAGCTCCACTTCGCGCGCATGATTTCAAGGATTCGATCTCAGATCT 696  
Db 85 TTGCTTGAAGTGTCTTCACTTTTCGGGTATGATTGAAGGATTCGGAATCTTCAAGTCT 144  
QY 697 CGTCTGTGACCGCGGTCAAGTCTTCAAGCTTCAAGCTTGAAGTGTCTCCGTTTTCAC 756  
Db 145 AGTCTGTTCGCGGACACATCTTCAAGCACTCAAGCTGGAAGTGTCTGATCTC 204  
QY 757 CACCGATGCTTTTCCATATGCTGCTTTTCAAGAGTTTAAAGTCTTGTTTTGA 816  
Db 205 CACCGATGCTCTACTATATGCTGCTTTTCAAGAGTCTTAAAGTCTTGTTTTGA 264  
QY 817 GGAAGCTCAATCTTGAAGACGAGATGCTTACAGAGTGTCTTGAATATAC 876  
Db 265 GGAAGCTCACTTTTGAAGATGATGCTGCTGCTCCATGAGTGTCTTGAATATAC 324  
QY 877 AGTCTTGAAGTCTCAATTTTCTTGAAGATGATGCTTGTGGAAGATTCAGGACCT 936  
Db 325 AGTCTTGAAGTCTCAATTTTCTTGAAGATGATGCTTGTGGAAGATTCAGGACCT 384  
QY 937 TGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 979  
Db 385 TGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 427

## RESULT 13

US-09-770-149-161  
; Sequence 161, Application US/09770149  
; Patent No. US20020059663A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Krickler, Maya  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil

APPLICANT: Hurban, Patrick  
TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
FILE REFERENCE: 2024 (PAPA-013PRV)  
CURRENT APPLICATION NUMBER: US/09/770,149  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,506  
NUMBER OF SEQ ID NOS: 999  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 161  
LENGTH: 719  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-770-149-161

Query Match  
Best Local Similarity 9.9%; Score 227.2; DB 9; Length 719;  
Matches 356; Conservative 0; Mismatches 173; Indels 5; Gaps 2;

QY 1593 GCTGTGACAAAGCTGAGGAGATTGCTCTATATCTCAGGCGTGGGGTTGACTGATGTAG 1652  
Db 11 GATGCAAGAACTCAGACGATTTGCATTCTATCTGAGACAAGGGCTTAAACGACTGG 70  
QY 1653 GCCTTGTTACATTGGACAATACAGTCCAAATGTGAGATGGATGCTGCTGGTTATGTGG 1712  
Db 71 GCTTAAGCTACATCGGACAGTACAGTCCAAACGTTGAGATGGATGCTGCTGGTTACGTAG 130  
QY 1713 GCGAGTCTGATGACGGCTTTGGAGTTGCTTAAGGGTCTCCTAGTCTTCAGAACTTG 1772  
Db 131 GTGAATCAGATGAAGGTTTAAATGGAATCTCAAGAGGCTGTCCAAATCTACAGAAGCTAG 190  
QY 1773 AAATGAGAGGTTGTTATTTTTCAGTCAAGCTGACATGCTGTGGCTGCAACAAATTGA 1832  
Db 191 AGATGAGAGGTTGT--TGCTTCAGTGGCGAGCAATCGCTGAGCGGTTTACAAATTCG 247  
QY 1833 CTCTCTTAGTACTTGTGGTGAAGGTTATGGTGTATCTCCATCTGGAAGCTGATCTTT 1892  
Db 248 CTTCACTGAGATACCTTGTGGTGAAGGTTACAGAGCATCGATGACGGGCAAGATCTAA 307  
QY 1893 TGGTAAGGCTGACCCCTTTTGGACATGATGTTGATCTCTTCCAGAAAGTGCTGCTACGA 1952  
Db 308 TGCAGATGGCTAGACCTGTCTGGAACATCGAGCTGATTCATCAAGAAAGATCCCGGAAG 367  
QY 1953 ATACCAATCCAGATGAGACTGTAGTTGTGAGCATCTGCTCATATTTCTTGCAATATT 2012  
Db 368 TGAATCAACAGGAGATGAAGAGATGGAGCATCCGGCTCATATATTGCTTACTACT 427  
QY 2013 CTCTTGAGGCGAGATCAGATTTTCCAGATCTGTTGCTGCTTGGACACTGCCACAT 2072  
Db 428 CTCTGGCTGGCCAGAACAGATTTGTCCAACTGTTAGAGTCTCTGAAGGAGCAATAT 487  
QY 2073 GCGTTGTATACCTAGAGCCAGAGCTGTGTATATATACAGTTTCTTTTGTGTTT 2126  
Db 488 GATATG--ACCCAAAACAGGTTTGTATATATAGATTTTGTAGTCTCGAGTTT 539

RESULT 14  
US-10-424-599-21331  
Sequence 21331, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 21331

LENGTH: 449  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_119267C.1  
US-10-424-599-21331

Query Match  
Best Local Similarity 9.6%; Score 220.2; DB 13; Length 449;  
Matches 228; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 552 CCGAGCCGCAATGTTCACTTATACCCGAGGATTGGGGCGGACACGTCACCTCCCTGGG 611  
Db 138 CTCAAGCCGCAATGTTTCACTTATACCCGAGGATTGGGGCGGACACATGTTAGTCCCTGGG 197  
QY 612 TCAAGAGATTCTCAGTACTTTCAGTTCGCTCAAGAGCTTCCACTTCCCGCGCATGATTG 671  
Db 198 TCAAGAGATTCTCAGTACTTTCAGTTCGCTCAAGAGCTTCCACTTCCCGCGCATGATTG 257  
QY 672 TCAAGATTCCGATCTTCAGAAATCTCGCTCGTGACCCGCGGTACGTCGCTTACGCTCTCA 731  
Db 258 TCAAGGATTCCGATCTTCAGAAATCTCGCTCGTGACCCGCGGTACGTCGCTTACGCTCTCA 317  
QY 732 AGCTTGACAAGTGTCTCCGGTTTCCACCACGATGCTCTTTTCCATATCGGTTCGCTTTTGA 791  
Db 318 AGCTTGACAAGTGTCTTCAGTTTCCACCACGATGCTCTTTTCCATATCGGTTCGCTTTTGA 377  
QY 792 A 792  
Db 378 A 378

RESULT 15  
US-10-424-599-21332  
Sequence 21332, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 21332  
LENGTH: 789  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_119268C.1  
US-10-424-599-21332

Query Match  
Best Local Similarity 9.6%; Score 218.6; DB 13; Length 789;  
Matches 227; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 552 CCGAGCCGCAATGTTCACTTATACCCGAGGATTGGGGCGGACACGTCACCTCCCTGGG 611  
Db 127 CTCAAGCCGCAATGTTTCACTTATACCCGAGGATTGGGGCGGACATGTTAGTCCCTGGG 186  
QY 612 TCAAGAGATTCTCAGTACTTTCAGTTCGCTCAAGAGCTTCCACTTCCCGCGCATGATTG 671  
Db 187 TCAAGAGATTCTCAGTACTTTCAGTTCGCTCAAGAGCTTCCACTTCCCGCGCATGATTG 246  
QY 672 TCAAGATTCCGATCTTCAGAAATCTCGCTCGTGACCCGCGGTACGTCGCTTACGCTCTCA 731  
Db 247 TCAAGGATTCCGATCTTCAGAAATCTCGCTCGTGACCCGCGGTACGTCGCTTACGCTCTCA 306  
QY 732 AGCTTGACAAGTGTCTCCGGTTTCCACCACGATGCTCTTTTCCATATCGGTTCGCTTTTGA 791

us-10-009-791-21.rnpb

Tue Apr 20 10:32:12 2004

Db 307 AGCTTGACAGTGCTTCAGTTTCACCACTGATGGTCCTTTCCACATCGGTGCTTTTGCA 366

Qy 792 A 792

Db 367 A 367

Search completed: April 20, 2004, 01:34:07

Job time : 1495 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 18:00:14 ; Search time 5811 Seconds  
(without alignments)

US-10-009-791-21  
11757.816 Million cell updates/sec

Title: US-10-009-791-21

Perfect score: 2288

Sequence: 1 gcacgagccacacgttaca.....acattcgaattttgtgaat 2288

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estm1:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	712	31.1	741	10	BE821935
2	614.2	26.8	2298	11	AY109490
3	598.2	26.1	2190	11	AY109565
4	564	24.7	588	14	CA820059

5	555.8	24.3	1811	14	CA654430
6	543.4	23.8	560	14	CF807278
7	542	23.7	574	13	BQ297870
8	538	23.5	538	13	BQ133928
9	532.2	23.3	561	12	BI967384
10	526	23.0	541	14	CF807291
11	515.6	22.5	538	12	BI974579
12	493.4	21.6	550	14	CA820349
13	489	21.4	493	9	AI855554
14	476.2	20.8	555	13	BQ630213
15	464	20.3	574	10	AW759107
16	462.6	20.2	589	10	BF325270
17	460.8	20.1	658	12	EG453637
18	451.6	19.7	783	14	CB894250
19	435.8	19.0	441	10	BE802940
20	417.8	18.3	455	12	BG156552
21	413.2	18.1	704	14	CB008682
22	391.2	17.1	630	14	CB001376
23	389.4	17.0	693	13	BU894657
24	385.6	16.9	465	10	AW396651
25	377.2	16.5	670	13	BU867812
26	375.2	16.4	542	13	BU080606
27	374.2	16.4	596	14	CB919625
28	373.8	16.3	512	10	AW185771
29	372.8	16.3	592	14	CB920883
30	372.2	16.3	442	10	BG045553
31	368	16.1	1683	11	AY109486
32	363.2	15.9	739	13	BQ997606
33	359.8	15.7	790	29	CG964083
34	359.6	15.7	382	10	AW595938
35	351.6	15.4	1134	14	CK209957
36	351.4	15.4	563	14	CB923093
37	351.4	15.4	813	28	BH678933
38	348.8	15.2	570	14	CB911803
39	341.4	14.9	556	12	BI893238
40	333.8	14.6	950	29	CG945532
41	333	14.6	652	13	BU548376
42	331	14.5	479	10	BG044386
43	330	14.4	428	14	CA936527
44	329.4	14.4	586	9	AI482980
45	326.8	14.3	588	10	AW222912

ALIGNMENTS

RESULT 1

BE821935/c

LOCUS

DEFINITION

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE821935 741 bp mRNA linear EST 24-MAY-2001  
GM700015B20H12 Gm-r1070 Glycine max cDNA clone Gm-r1070-6144 3',  
mRNA sequence.

BE821935

EST.

Glycine max (soybean)

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 741)

Vodkin, L., Kelm, P., Shoemaker, R., Retzel, E., Khanna, A., Corvett, V.,

Erpelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.

A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)

Other ESTs: AW185771 corresponding to Gm-cl019-1292 (5')

Contact: Vodkin, L.O., PI, A Functional Genomics Program for

Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582  
 Email: 1-vokine@uic.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134. For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 913-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or info@genome  
 systems.com web site: www.genomesystems.com  
 Seq primer: 5'-TTTTTTTTTTTTTTTTT (A/C/G)-3'.

## FEATURES

location/Qualifiers  
 1. 741  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="Gm-r1070-6144"  
 /clone\_lib="Gm-r1070"  
 /note="The library Gm-r1070 is a sequence-driven, reracked  
 set of 9,216 clones selected from cDNA libraries from  
 various tissues and stages of development of soybean that  
 represent 2,639 sequences from immature cotyledons, 1,770  
 from immature seed coats, 3,938 from flowers, and 869  
 from young pods. The 5' ESTs of the source clones from  
 the different libraries was used to select singletons, or  
 a representative of each contig, which were reracked to  
 form library Gm-r1070. The cDNA clones of the reracked  
 Gm-r1070 library were then sequenced at the 3' end. The  
 contig analysis to select unique genes was performed by  
 the laboratory of Ernest Retzel, Center for Computational  
 Genomics and Bioinformatics, University of Minnesota.  
 http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html  
 . Reracking was performed by Genome Systems, St. Louis,  
 http://www.genomesystems.com, and 3' sequencing by the  
 Keck Center for Comparative and Functional Genomics,  
 University of Illinois.  
 http://www.life.uiuc.edu/biotech/keck.html. Note: The  
 corresponding 5' EST from each clone in the Gm-r1070  
 library is listed in the 'OTHER EST' field. The detailed  
 information on the source library for each clone can also  
 be obtained by referring to the Genome Systems clone ID of  
 the original cDNA library that is also listed under  
 'OTHER EST'."

## ORIGIN

Query Match 31.1%; Score 712; DB 10; Length 741;  
 Best Local Similarity 96.1%; Pred. No. 7.2e-123;  
 Matches 712; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

1523 GTGTTGCTGACCATGAAGAAGATACATGTTGCCACTGCAATGGGGTGGGCT 1582  
 741 GTGNNCNCNNNCCATGAAGAAGATNCTGATTTGCCACTGNNNAATGGGNTGAGGNNN 682

1593 CTACTGAGGGCTGTGCAAGCTGAGGAGATTGCTCTATATCTCAGCGTGGCGGGTTG 1642  
 681 CTACTGAGGGCTGTGCAAGCTGAGGAGATTGNNCTATANNCCAGCGTNNCGGGTTG 622

1643 ACTGATGTAGGCTTGGTTTACATTCGCAATACAGTCCAAATGTGAGATGATGCTTT 1702  
 621 ACTGATGTANNNNNTGGTTTACATTGNACATACAGTCCAAATGTGAGATGATGCTTT 562

1703 GGTATGTGGGGAGCTGTGATCAGGGCTTTGGAGTTGCTTAAGGGGTGCTAGTCTTT 1762  
 561 GGTATGTGGGGAGCTGTGATCAGGGCTTTGGAGTTGCTTAAGGGGTGCTAGTCTTT 502

1763 CAGAACTTGAATGAGAGGGTGTATTTATTTTTCAGTGAACGTGCACTTGCTGTGGCTGCA 1822  
 501 CAGAACTTGAATGAGAGGGTGTATTTTTCAGTGAACGTGCACTTGCTGTGGCTGCA 442

1823 ACACAAATGACTTCTCTTAGGTACTTGTGGTGCAGAGTTATGTTGTTATCTCCATCTGA 1882  
 441 ACACAAATGACTTCTCTTAGGTACTTGTGGTGCAGAGTTATGTTGTTATCTCCATCTGA 382

1883 CGTGATCTTTTGGTAATGCTCGACCCCTTTTGGACATTCAGTTGATCTCTCTAGAAAG 1942  
 381 CGTGATCTTTTGGTAATGCTCGACCCCTTTTGGACATTCAGTTGATCTCTCTAGAAAG 322

1943 GTGGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTGAGCATCTGCTCATATCTTT 2002  
 321 GTGGCTACGAATACCAATCCAGATGAGACTGTAGTTGTTGAGCATCTGCTCATATCTTT 262  
 2003 GCATATATTCTCTTCCAGGGCAGAGATCAGATTTTCCAGATACTGTTGTGCGCTTGGAC 2062  
 261 GCATATATTCTCTTCCAGGGCAGAGATCAGATTTTCCAGATACTGTTGTGCGCTTGGAC 202  
 2063 ACTGCCACATCGCTTGATACCTAGAGCCAGAGCTGTGTATATATACACGTTTCTTTTG 2122  
 201 ACTGCCACATCGCTTGATACCTAGAGCCAGAGCTGTGTATATATACACGTTTCTTTTG 142  
 2123 TTTTCTTCTCCCTTTTCATATGCTGTTCTATGTTCCCTCTCTATTTGATGTCATTTT 2182  
 141 TTTTCTTCTCCCTTTTCATATGCTGTTCTATGTTCCCTCTCTATTTGATGTCATTTT 82  
 2183 AGACAATAGTCTGTATTAAGCCTGTGTTTTCATTTGAAATCTGAAACGCTTCCCTTA 2242  
 81 AGACAATAGTCTGTATTAAGCCTGTGTTTTCATTTGAAATCTGAAACGCTTCCCTTA 22  
 2243 ACCGTATTGGCTCCCTTAAAA 2263  
 21 ACCGTATTGGCTCCCTTAAAA 1

RESULT 2  
 AV109490  
 LOCUS AV109490 2298 bp mRNA linear HTC 17-OCT-2002  
 DEFINITION Zea mays CL920\_1 mRNA sequence.  
 ACCESSION AY109490  
 VERSION AY109490.1 GI:21213230  
 KEYWORDS HTC.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 2298)  
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
 Zea Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 REFERENCE 2 (bases 1 to 2298)  
 Coe, E.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
 COMMENT If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
 www.zmdb.iastate.edu.  
 FEATURES  
 Location/Qualifiers  
 1. 2298  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:633078"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed DuPont contigs; this resource was  
 assembled by DuPont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

ORIGIN  
 Query Match 26.8%; Score 614.2; DB 11; Length 2298;



Best Local Similarity 59.9%; Pred. No. 1.4e-104; Matches 1037; Conservative 0; Mismatches 680; Indels 15; Gaps 4;			
QY	345	GTGTGGTCCGAGTGGTCTCGTACGTCGGTCACTCCCTTACATCGACGACCCCAAGGACCGCG	404
DB	275	GGTTCGCCGAGGAGATGCTGCACTCGTATGGGTTCGTGCGAGNNNNNNNNNNNN	334
QY	405	ACGCCGTTTCCAGGTGTGTCAGTGGTTCAGAGCTCGATCGTCACTCCGCGACGACG	464
DB	335	NNNNNNNTGCGTGTGTGTCAACCGGTGGCAACCGGTGCGACGCGCTCTCGCGAAGCACG	394
QY	465	TCACCATCGGCTCTGCTACACCAACCGCGGTCTCGCGCGCGCTTCCCGCACCC	524
DB	395	TGACGGTCCCTTCTGCTACGCGGTTTCCCGGACGCTGCTCGCGGGTTCGCGCGG	454
QY	525	TGAGTTCGCTCAAGCTCAAGGCGCAAGCCCGAGCGCAATGTTCAACTTGTATACCCGAGG	584
DB	455	TGAGTTCGCTCGGCTGAAGGGAAGCCCGCGCGCCATGTACGGGCTCATACCCGACG	514
QY	585	ATTGGGGCGGACAGTCACTCCCTGGGTCAAGAGATTTCTCAGTACTTCCGATTCCTCA	644
DB	515	ACTGGGGCGCTTACCGCGCGCGTGGATCACCGAGCTCGCGCGCGCTCGAGTGCCTCA	574
QY	645	AGAGCTCCACTTCCGCGCGCATGATTGTCAAGGATTCGATCTTCAGAACTCTCGCTCGT	704
DB	575	AGGCGCTCCACTCCGACGCGATGCTGCTCAAGAGCGACCTCGCGCGAGCTCTCGCTG	634
QY	705	ACCGCGCTCAGTCTTCAAGCTTCAAGCTTGAACAAGTGTCTCGGTTTTCACCAAGGAT	764
DB	635	NN	694
QY	765	GTCTTTTCATATCGTTCGCTTTTCCAGAGTTTAAGAGTCTTGTTTTGGAGGAAAGCT	824
DB	695	GACTCCGCTCTGCTGCGCGCTCCTCGAGATCACTGAGGACTTANNNNNNGAAGATGTC	754
QY	825	CAATTTCTGAGAGGACGAGAAATGGCTACAGAGCTTCTCTTGAATATACAGTTCTTG	884
DB	755	AAATTGATGATAGGCGAGTGAATGATCCAGATCTCGAGTCTGCTGCTGCTGTTCTGA	814
QY	885	AGACTCTCAATTTTACTTTGACAGACATTTGCTGTTGTAAGATTTGAGGACCTTGAACTTT	944
DB	815	CAACATTTGAATTTCCACATGACTGAGCTTGA--GTGATGCGAGCTGACCTTAAAGCTTC	871
QY	945	TAGCTAAATTTGCCCAACTAGTCTCTGTAACCTTACTGACTGTGAATTACTGGATC	1004
DB	872	TTGCAAGAGCTGCAAGTCACTGATTTTCAATGAAGATTTAGTACGCTCTTTGATTT	931
QY	1005	TTGTGAATCTTTTAAGCATGCTCTGCGCTGGAAGAGTTTGTGGAGGACCTTACAACG	1064
DB	932	TGATAGATTTCTCCAAATTTGCCACAGCACTGGAAGAAATTTGCTGGAGGACATTCATG	991
QY	1065	AGAACCCAGAA-----AGATCTCTGCTATATCATTAACAGCAAGTTATGTCGATGG	1118
DB	992	AGCAAGGGGAACTCAGCAAGTATGTGAATGTTAAATTTCCATCAAGACTATGCTCTCG	1051
QY	1119	GTTTACATATTTGGAAGAAATGAGTGTGCCATTTGTTTCATGTTTGGAGCGGTACTAA	1178
DB	1052	GACTTACTTACATGGGAAACAAATGAATGCCCATTTATGTTCCNNNNNNCTGCATACTAA	1111
QY	1179	AAAAATTGATCTCTCTATGCAATGCTAGACACGAGGATCATTTGTATGTTAAATCAAA	1238
DB	1112	AGAGCTGGATTTGCAATACACTTTCTCCACCACTGAGGACCATTTGCCAGCTCATTGCAA	1171
QY	1239	GGTGTCCAAATCTGGAAGTCTTTGAGACAGGATGTAAATTGGAGATAGGGTTAGAGG	1298
DB	1172	AATGCCCGAACTTACTAGTTCTCGCGGTGAGGAATGTGATTGGAGATAGAGGATTAGGAG	1231
QY	1299	TTCTTGTCTGTGTGTAAGAGGCTTAAAGGCTTAGGATTTGAAAGGGCGGATGATGATC	1358
DB	1232	TTGTTGGGATAGCTGCAAGAGCTTCCAAAGGCTCAGATAGAGCGAGGATGATGAAG	1291
QY	1359	AAGGAATGGAGGATGAAGAAGGTACTGTGTCCCATAGAGGGCTTAATAGCCCTTGTCAAGG	1418

Db	1292	GAGCTGTGCAAGAGAGCAGGGAGGGGTCTCTCAAGTGGGCTTGACGGCTATAGCCGTAG	1351
QY	1419	GCTGTTCAAGAGCTTGAATACATGCTGTTTATGTTGCTGATATACAAATCATCTCTGG	1478
DB	1352	GTTCCCGTGAAGCTGGAATATATAGCTGCTATGTTGCTGATATACCAATGGGCGCTGG	1411
QY	1479	AACATATTGGAACCTCACTTGAAGAACCTCTGCTGATTTTTCGCTTGTGTTGTTGACCATG	1538
DB	1412	AATCTATCGGCACATTTCTGNNNNNNCTATACGACTTCCGGCTTGTCTACTTGATAGAG	1471
QY	1539	AAGAGAAGATAATGATTGGCCACTTGACAATGGGCTGAGGGCTCTACTGAGGGGCTG	1598
DB	1472	AAGAGAGGATAACAGACTTGGCACTGGACAATGGTGTCCGAGCTTGTGTGAGGGGCTGCA	1531
QY	1599	ACAAGCTGAGAGATTTGCTCTATATCTCAGCGGTGGCGGTTGACTGATGTAGGCGCTTG	1658
DB	1532	CCAAGCTTCGGAGGTTTGTCTCTGTAATTGAGACAGGAGGCTCTCAGATCAGCTCTCG	1591
QY	1659	GTTACATTTGGAACAATACAGTCCAAATGTGAGATGGATGCTGCTTGTGTTATGTGGGGAGT	1718
DB	1592	GCTACATTTGGACAGTGCAGCGGAAACATCCAGTACATGCTTCTCGTAAATGTGGGAAA	1651
QY	1719	CTGATGCAAGGCTTTTGGAGTTTCGTAAGGGGTGCTCTAGTCTTCAGAACTTGAATGA	1778
DB	1652	CTGATGATGATGATGATCAGCTTCGCAATGGGTTGCGTAAACCTGCGAAAGCTTGAACCTCA	1711
QY	1779	GAGGCTGTTTATTTTTCAGTGAACCTGACACTTGTCTGCTGCTGCAACACAAATGACTTCTC	1838
DB	1712	---GAGTTGCTGCTTTCAGCGAGCGACCTGGCCCTTGCATATACATATGCTTCTCC	1768
QY	1839	TTAGTACTTGTGGGTGCAAGTTAAGTGTATCTCCATCTGCGAGCTGATCTTTTGGTAA	1898
DB	1769	TGAGTACTGATGCTTTCAGGCTTACAAAGGCTCTCAAAACCGCGGAGACCTCATGCTCA	1828
QY	1899	TGGCTCGACCTTTTGGACATTTGAGTTGATTTCTTC---TAGAAGGTGGCTACGAATA	1955
DB	1829	TGGCAAGGCTTCTTGGAAACATAGAGTTTACACCTCCCAATCCTAAGAACGGAGTTGCG	1888
QY	1956	CCAATCCAGATCAGACTGTAGTTTGTAGCATCTGCTCATATTTCTTGATATTTATCTC	2015
DB	1889	TGATGGAAGTGGGAGGCTTGTGTAGATGTCAGCTCAGATCTTGCATACCACTCC	1948
QY	2016	TTGCAAGGAGATGATGATTTTCCAGATCTGTTGCTGCTTGGCTTGGACATGC	2067
DB	1949	TGCGCGTAAAGAGGCTGGAGTCCCAATCCGCTGCTTCTTGTATCCTGC	2000

RESULT 3			
LOCUS	AY109565	2190 bp	mRNA linear HTC 17-OCT-2002
DEFINITION	Zea mays CL896_1 mRNA sequence.		
ACCESSION	AY109565		
VERSION	AY109565.1	GI:21213333	
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
AUTHORS	Coe,E.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,		

www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers  
1..2190  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:633027"  
/clone\_lib="Maize Mapping Project/DuPont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed DuPont contigs; this resource was  
assembled by DuPont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

ORIGIN

Query Match 26.1%; Score 598.2; DB 11; Length 2190;  
Best Local Similarity 60.6%; Pred. No. 1.3e-101;  
Matches 1053; Conservative 0; Mismatches 665; Indels 21; Gaps 6;  
345 GTGTGGTCCAGGTGTCCTCGACTGCGTTCATCCCTTACATCGACGACCCCAAGGACCGCG 404  
231 GCCTCCCGGAGGCGCTGCACCTGTGTTCGGCTACATGAGGACCCCGGAGACCGCG 290  
405 ACGCCGTTTCCAGGTGTGTCGACGCTGTCGATCGAGTTCGATCGCTCAACCGCAAGCAG 464  
291 AGCGGCTCGCTGTCGTCGCGCTCTGCGACCGCATCGACGGCTCTCGCGAAGCAG 350  
465 TCACCATGCGGCTCTGCTACACACACCGCGCTCGCTCGCGCGCTTCCCGCACC 524  
351 TCACCGTGGGCTTCTGCTACCGCGTGGAGCCGCGCGGCTGCTCGCCAGGTTCCCGGGC 410  
525 TCAGTGCCTCAAGTCAAGGCGAAGCCCGGACCGCAATGTTCAATTCGATACCGGAG 584  
411 TCAGTGCCTCGCGCTCAAGGCGGAGCCCGCGCGCCATGATAGCGGCTCATCCCGAAG 470  
585 ATTGGGCGGACACGTCACCTCCCTGGGTCAAGAGATTTCTCAGTATCTCGATGCTCTCA 644  
471 ACTTTGGCGCTACGCGCGGCTCGGTTCGCGGAGCTCGCGCGCTTCTGCTGCTCA 530  
645 AGAGCTTCACCTCCGCGGCTATGTCAGGATTCGATCTTCAGATCTCGCTCGTG 704  
531 AGGCGCTCCACTCGCTGCGATGACCGTCAACGAGGACGATCGCGTGTCTCTCCAG 590  
705 ACCGCGTCACTGCTTCCAGCTCTCAAGCTTGACAGTGTCTCGGTTTCAACCGATG 764  
591 CGCGGCGCACATGCTACGCGGCTCAAGCTCGACAGTGTCTCGGCTTCTCAACAGNN 650  
765 GTCTTTTCCATATCGGTCGCTTTTGCAAGATTTAAAGATCTTGTGTTGAGGAAAGCT 824  
651 NNN 710  
825 CAAATCTTGAGNAGACGAGATGCTGCTACGAGCTTGTGTTGAATTAATACAGTTCTTG 884  
711 TAAATGAGATGAAGGAGTGAATGGCTCCATGAATCGCTGTCAACATTTCTGTTCTGG 770  
885 AGACTCTCAATTTTACITGACAGACATGCTGTTGTGAGATTTGAGACCTTTGAATTTT 944  
771 TGACACTGAATTTTACATGACAGAACT---CAAAGTGGAGCTGCTGATCTGGAGCTTC 827  
945 TAGCTAAATAATGGCCCAACTTAGTGTCTGTAACCTTACTGACTGTGAAATCTGGATC 1004  
828 TTGCAAGAAGACTGTAATCAITTAATTTCTTGAAGATGGGTGACTGTGATCTTTCAGATC 887  
1005 TTGTGAACTCTTTAGCATGCTCTGCTGCTGGAAGATTTTGTGGAGCACTTCAACCG 1064  
888 TGATTTGGNNNNNNCCAAACCTTCAAGCAATTCGAAGATTCGCGGAGGTGCGNNNNNG 947  
1065 AGGAACACGA-----AAGACTCTGCTATATCATTACCAGCAAGTTATGTCGATT--- 1116

Db 948 AAGTTGAGAGTACACCAAGTACGAAAGGTCATTTTCCACCTAGGCTATGCTTCTTAN 1007  
QY 1117 -GGGTTTAAACATATATTGGAAGAATGAGTTGCCATTTGTTTCTATGTTTCAGCGGTAC 1175  
Db 1008 NNNNTCTTACCTTCATGGTNNNNNCGAGATGCCAGTTATCTTTCCGTTATTCACAAATGC 1067  
QY 1176 TAAAAAATTGGATCTCTCTATGCAATGCTAGACAGGAGATCATATGTTTAAATCC 1235  
Db 1068 TTAAGAACTGGACTTGCAGTTCTTCCCTCACCACTGAAGACCACTGTGAGCTTATG 1127  
QY 1236 AAGGTGTCCAAATCTGGAAGTCTTGGAGACAAGGAATGTAATGGAGATAGAGGTTAG 1295  
Db 1128 CTAATGCCCCAACCTTAAGTGTCTCGAGGTGAGGATGTGATTGGGATAGAGACTAG 1187  
QY 1296 AGTTCTTGGTGTGTTGTTAAGAGGCTTAAAGGCTTAGGATTTGAAAGGGCGGATGATG 1355  
Db 1188 AAGTTGTTGCTGTACATGCAAGAGCTACGAAGGCTCAGAAATGAGGCTGGCGAGATG 1247  
QY 1356 ATCAAGGAATGGAGATGAAGAGGTACTGTGTCCCATAGAGGGCTAATAGCTTGTCTAC 1415  
Db 1248 ATCTTGGCCAAAGAGAGAGAGGAGG---AGTCTCTCAGATAGTTTGACGCGTAGCG 1304  
QY 1416 AGGCTGTTCAGAGCTTGAATACATGCTGCTGTTTATGTCTCATATTAACAATGATCTC 1475  
Db 1305 TTGGCTGCGCGAACTAGAGTACATAGCTGCTGCTATGTATCTGATATCAAAATGGGCGC 1364  
QY 1476 TGGAAATATTTGGAACTCACTTGAAGAACTCTGTGATTTTCCGCTTGTCTTACTTGACA 1535  
Db 1365 TGGAAATCCATTGGCACTTCTGCAAGATCTATATGATTTCCGCTTGTCTTACTTGACA 1424  
QY 1536 ATGAAGAGAGATTAAGTATTTGCCATTTGCAATGCGGTGAGGGCTCTACTGAGGGGCT 1595  
Db 1425 AACAGAAAGAGATAACAGACCTGCCACTTGAACAGCGGTCCGCTGCTTACTGAGAACT 1484  
QY 1596 GTGACAAGCTGAGGAGATTTGCTCTATATCTCAAGCGTGGCGGGTGTGACGTGATGAGGCC 1655  
Db 1485 GGTCAAACTTCGAGGTTTGCCTTCTACTGAGACGGGAGGGCTTTCAGACGTAGGCC 1544  
QY 1656 TTGGTTACATTTGACAATACAGTCCAAATGTGAGATGGATGCTGCTTGGTTATGTTGGGG 1715  
Db 1545 TCGTTTACATTTGACTATACAGTGGGAACATCCAAATACATGCTGCTCGGCAAGCTTGCTG 1604  
QY 1716 AGTCTGATGACGGCTTTGGAGTTCGTAAGGGGTGCTCTAGTCTTTCAGAAACTTGAAA 1775  
Db 1605 AATCTGACAAAGGATGATCCAGTTTGAACGGGATGCAACCACTTCGCGAAGCTCGAGC 1664  
QY 1776 TGAGAGGCTGTTTATTTTTCAGTGAACGTGACCTTGTGCTGGGTGCAACCAATGACTT 1835  
Db 1665 TGAGGGGCTG---CTGCTTTCAGCGAGCGAGCTCTGGCAGTGGCGCTGCTCCAGATGCCCT 1721  
QY 1836 CTCTTAGGTACTGTGGGTGCAAGGTTATGTTGATCTCTCCATCTGGACGATGATCTTTGG 1895  
Db 1722 CGCTGAGGTACATATGGGTGCGAGGGGTACAGAGCTCTCGAAACGGGCGAGGACCTCATGC 1781  
QY 1896 TAATGGCTCGACCTTTTGGAACTTTCAGTT---GATTCTCTTCTAGAAAGGTGCTCAGCA 1952  
Db 1782 TCATGGCCAGGCCGTACTGGAACTTGAATTCGCGCTCCCATTCCTCCGAGAGCGCTTATC 1841  
QY 1953 ATACCAATCCAGATGAGACTGTAGTTGTGAGCATCTCTGCTCATATTTCTTGCATATATT 2012  
Db 1842 GGGTGTGGCAGATGGACAGCTTGTGTTGACACCCATGCCAGGTCTCTCGGTATTACT 1901  
QY 2013 CTCTTGACGGCCAGAGATCAGATTTTTCAGATACTGTTGCTGCTTGGACACTGCCACA 2071  
Db 1902 CCCCTGCTGGAAGAGGCGCGACTGCTCCTAGTGGGTGAGCTTTGACCTTGGACCTTGGTCA 1960

RESULT 4

CA820059

LOCUS

DEFINITION

CA820059 588 bp mRNA linear EST 09-DEC-2002  
sau83hl1.1.y1 Gm-cl048 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl048-5685 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.



Matches 1000; Conservative 0; Mismatches 574; Indels 21; Gaps 8;			
QY	353	GACGTGGTCTCGACTCGGTCTATCCCTTACATCGACGACCCCAAGGACCGGACCGCCGTT	412
Db	70	GACGTGGGCTGGGGCTGGTCAATGGGTGCGTGGAGGACCCCTGGGACCGGACCGCCATC	129
QY	413	TCCAGGTGTGTCGACGCTGGTACGAGCTCGACTCGCTACCCCGAAGCAGCTACCAATC	472
Db	130	TGCTCTGGTCTGGCGCACTGGTGCAAGGTGACGGCTCGACGGCGACGACGACGACGTC	189
QY	473	GCGCTCTGTACACACACACCCCGGCTCGCTCCGCGCGCTTCCCGACCTCGAGTCG	532
Db	190	GCCATGGCTACTCTCCACACCCCGGACCGCTCTTCCGCGGCTTCCCTGCTCGAGTCG	249
QY	533	CTCAGCTCAAGGGCAAG-CCCGAGCGGCATGTTCACTTGATACCCGA-GGATTGGG	590
Db	250	CTCAAGCTCAAGGCCCAAGCCCCCGGCTCCATGTTCAACCTCATCCCGAGGAGTGGG	309
QY	591	GCGGACAGCTACTCCCTGGGTCAAGAGATTCTCAGTACTTCGATTGCTTCAAGAGCC	650
Db	310	GGAGCTCGCTCGCCCTGGATCGCGAGCTCTCCGCTTCTTCCACTTCTTCAAGATGC	369
QY	651	TCCACTTCGCGCGCATGATGTCAAGGATTCGATCTTCGAAATCTCGCTCGTGACCGG	710
Db	370	TGCACCTCGCGCGGATTATTGCTCCGACCAACCTCCCGCTGCTCGTGGCGGCAAGG	429
QY	711	GTACGTGTTTCAAGTCTCAAG-CTTGACAGTCTCGGTTTCAACACCGATGGTCTT	769
Db	430	GCCAAATGCTGCTTCCCTCGAGCCCTGGACCGCTCGCTCGGTTTCCAACTCTTCCCTC	489
QY	770	TTCCATATCGGTCTGTTTCAAGAGTTTAAAGAGTCTTCTTTTGGAGGAAAGCTCAATT	829
Db	490	GGTCTCTCGCGGTGNTGGAGAACTGGAAACGTTCTTCTTGAAGAAAGTCTGTT	549
QY	830	CTTGAAGG---ACGGAGAAATGGCTPACAGAGCTTGTCTTGAATTAACAGTCTTGAG	886
Db	550	GCTGAGAAAGAAATGATGAATGGCTCCGTGAGCTTGTACCAAGCAATATCTGTCTTGAG	609
QY	887	ACTCTCAATTTTCTGTGACAGACATGCTGTGTGAAGATTGAGAGACCTTGAATTTTA	946
Db	610	AGCTGAATTTCTTCTGACGGA---TCTCAGGGCAATCCCTGCAATCTTCTCTCCCT	666
QY	947	GCTAAATTTGCCCAACTTAGTGTCTGTGAACTTACTGACTGTGAAATATCTGATCTT	1006
Db	567	GTGCGAAATTTGCCAAGGCTGAAACTCTCAAGATTAGCGACTGTTTCAATGCTGACCTG	726
QY	1007	GTGAATCTTTAAGCATGCTCTGCGCTGGAGAGTTTGTGGAGGACCT-----AC	1060
Db	727	GTGACCTGTTCGTACAGCAGAAACACTACAGACTTTTGTGCTGGTTCCTTTGATGAT	786
QY	1061	AACGAGAACCAAGATACCTGCTATATCATTACAGCAAGTATGTCGATTGGGT	1120
Db	787	CAAGATCAAGGTGGGAATTATGTAATACTATTTCCTCTTCACTACAGCACTTGAT	846
QY	1121	TTAACAATATTTGAAAGATAGTTGCCCCATTTGTTCATGTTTGCAGCCGCTACTAAA	1180
Db	847	TTGCTCTACATGGAAACAAATGAGATGACATATTTTCCATATGTTGTCACACTCAAG	906
QY	1181	AAATTGGATCTCTTATGCAATGCTAGACACGAGATCAATTTGATTTTAAATCCAAAGG	1240
Db	907	AAGTTGGACCTTCAGTTTCACTTCTTACCACAGAGATCACTGTCAATTAGTCCAGCGC	966
QY	1241	TGTCCAATCTGAGAGCTCTTGACAGCAAGAAATGTAATTTGAGATAGAGGTTTAGAGTT	1300
Db	967	TGCCCAATCTAGAATTTTGGAGGTGAGGATGTAAGAGATCGAGGTTTAGAGTT	1026
QY	1301	CTTGGTCTGTTTGAAGAGCTTAAAGAGGCTTAGGATTTGAAGGGCGGATGATGATCAA	1360
Db	1027	ATTGCGGGACCTGCAGAAATATACGCACTCAGAGTTCAGAGAGAGAGATGATGACCAA	1086
QY	1361	GGA---ATGAGAGATGAAGAGTACTGTGTCCCATAGAGGGCTAATAGCTTGTACAG	1417
Db	1087	GGAGGTCTTGAGGACGAACAGGGTAGAGTGAACACAGTAGGATTTGATGGCTGTAGTGAA	1146

## RESULT 6

CF807278/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CF807278 560 bp mRNA linear EST 27-OCT-2003  
PSHB022K06f USDA-IPAFS:Expression of Phytophthora sojae genes  
during infection and propagation Phytophthora sojae cDNA clone  
SHB022K06 5, mRNA sequence.

CF807278

CF807278.1 GI:37995689

EST

Phytophthora sojae

Phytophthora sojae

Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

Phytophthora.

1 (bases 1 to 560)

Tyler, B.

Tyler, B. Not Published

Unpublished (2003)

Contact: Tyler B

Tyler lab

VBI

1880 Pratt Dr., Blacksburg, VA 24061, USA

Tel: 540-231-7318

Email: bmtylev@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 022 row: K column: 06

Seq primer: BK reverse primer

High quality sequence stop: 560.

Location/Qualifiers

1. .560

/organism="Phytophthora sojae"

/mol\_type="mRNA"

/db\_xref="taxon:67593"

/clone="SHB022K06"



1837 TCTTAGGTAAGTGGGTCGAAGGTATAGGTATCTCCATCTGACGTGATCTTTTGGT 1896  
 361 TCTTAGGTAAGTGGGTCGAAGGTATAGGTATCTCCATCTGACGTGATCTTTTGGC 420  
 1897 AATGGCTGACCCCTTTGGAGCAATGAGTATCTCTCTAGAGAGGTGGCTACGAATAC 1956  
 421 AATGGCTGACCCCTTTGGAGCAATGAGTATCTCTCTAGAGAGGTGGCTACGAATAC 480  
 1957 CAATCCAGATGAGCTAGTGTGTGAGCATCTCTCATATTTCTGATATTAATCTCT 2016  
 481 CAATTCAGATGAGCGGTAGTGTGAGCATCTCTCATATTTCTGATATTAATCTCT 540  
 2017 TCCAGGGCAGAGATCAGATTTTCCAGATCTGTT 2050  
 541 TGCAGGGCAGAGATCAGATTTTCCAGATCTGTT 574

RESULT 8  
 BQ133928  
 LOCUS  
 DEFINITION  
 san55a01.y1 Gm-cl052 Glycine max cDNA clone SOYBEAN CLONE ID:  
 Gm-cl052-3458 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.  
 ; mRNA sequence.  
 BQ133928  
 BQ133928.1 GI:20207839  
 EST.  
 Glycine max (soybean)  
 ORGANISM  
 Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 538)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
 Bowers, J., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estowatson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact: ccu@resgen.com web site:  
 www.resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 421.  
 Location/Qualifiers  
 1..538  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl052-3458"  
 /tissue\_types="whole seedlings of greenhouse grown plants"  
 /dev\_stage="1 week old"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl052"  
 /note="vector: pbluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The Harcoy N1L was constructed and seed was  
 provided by Dr. J. Specht, University of Nebraska  
 (Shoemaker and Specht, 1995). The cDNA library was  
 constructed from mRNA isolated from whole seedlings of 1  
 week old greenhouse grown plants. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with a XhoI restriction site and a 3'  
 anchor. EcoRI adapters were ligated to the blunt-ended

cDNA fragments followed by XhoI digestion. The cDNA  
 fragments were directionally cloned into the EcoRI-XhoI  
 restriction site of the pBluescript vector. The ligated  
 cDNA fragments were transformed into DH10B host cells  
 (GibcoBRL). The library was constructed in cooperation  
 with Dr. Paul Keim's laboratory at Northern Arizona  
 University."

ORIGIN  
 Query Match 23.5%; Score 538; DB 13; Length 538;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-90;  
 Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 307 ATCCAAATATGACGGAGGAACGGAGCTGCGGAGACACGTGTGTGTCGAGTCTCTCGA 366  
 DB 1 ATCCAAATATGACGGAGGAACGGAGCTGCGGAGACACGTGTGTGTCGAGTCTCTCGA 60  
 QY 367 CTGCGTCAATCCCTTACATCGACGACCCCAAGACCGCGCGTTTCCAGGTGTGCG 426  
 DB 61 CTGCGTCAATCCCTTACATCGACGACCCCAAGACCGCGCGTTTCCAGGTGTGCG 120  
 QY 427 AGCTGTGTACGAGCTGCGTCTCACCGGACGACGTCACCATCGGCTCTGTACAC 486  
 DB 121 ACGCTGTGTACGAGCTGCGTCTCACCGGACGACGTCACCATCGGCTCTGTACAC 180  
 QY 487 CACCAACCCCGGCTCGCTCCGCGCGCTTCCCGCACCTCGAGTCGCTCAAGCTCAAGGG 546  
 DB 181 CACCAACCCCGGCTCGCTCCGCGCGCTTCCCGCACCTCGAGTCGCTCAAGCTCAAGGG 240  
 QY 547 CAAAGCCCGGAGCGCAATGTTCAACTTGTATACCCGAGGATTTGGGCGGACAGTCACTCC 606  
 DB 241 CAAAGCCCGGAGCGCAATGTTCAACTTGTATACCCGAGGATTTGGGCGGACAGTCACTCC 300  
 QY 607 CTGGGTCAAAGAGATTTCTCAGTACTTTCAGTTCCTCAAGAGCTTCCACTTCGCGCAT 666  
 DB 301 CTGGGTCAAAGAGATTTCTCAGTACTTTCAGTTCCTCAAGAGCTTCCACTTCGCGCAT 360  
 QY 667 GATTGTCAAGGATTCGCGATCTTTCAGATCTCCCTGTGTACCCGCGGTTCAGTCTTCAAGC 726  
 DB 361 GATTGTCAAGGATTCGCGATCTTTCAGATCTCCCTGTGTACCCGCGGTTCAGTCTTCAAGC 420  
 QY 727 TCTCAAGCTTGACAAAGTGTCTCCGGTTTACCAACCGAGTCTTTTCCATATCGGTGCTT 786  
 DB 421 TCTCAAGCTTGACAAAGTGTCTCCGGTTTACCAACCGAGTCTTTTCCATATCGGTGCTT 480  
 QY 787 TTGCAAGATTTAAGAGTCTTCTTTTTCGAGGAAGTCAATTTCTTGAGAAGACGGA 844  
 DB 481 TTGCAAGATTTAAGAGTCTTCTTTTTCGAGGAAGTCAATTTCTTGAGAAGACGGA 538

RESULT 9  
 B1967384/c  
 LOCUS  
 DEFINITION  
 Gm830001B20C10 Gm-r1083 Glycine max cDNA clone Gm-r1083-140 3',  
 mRNA sequence.  
 B1967384  
 B1967384.1 GI:16341789  
 EST.  
 Glycine max (soybean)  
 ORGANISM  
 Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 561)  
 Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,  
 Erpelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.  
 A Functional Genomics Program for Soybean (NSF 9872565)  
 Unpublished (1999)  
 Other ESTs: AW102154 corresponding to Gm-cl009-1054 (5')  
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
 Soybean (NSF 9872565)  
 Lewin, H. A., Director, Keck Center for Comparative and Functional



Genomics	University of Illinois	Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA		
Tel:	(217) 244-6147			
Fax:	(217) 333-4582			
Email:	l-vodkin@uiuc.edu			
This clone is available through:	Incyte Genomics, 4633 World Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or (314) 427-3222 FAX: (314) 427-3324. Web site: http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio	n/index		
Seq primer:	5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.			
Location/Qualifiers				
1..561				
/organism="Glycine max"				
/mol_type="mRNA"				
/db_xref="taxon:3847"				
/clone="Gm-r1083-140"				
/clone_lib="Gm-r1083"				
/note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Supernod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nfsioy/. Reracking was performed by Incyte Genomics, St. Louis, http://www.incyte.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html. Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."				
ORIGIN				
Query Match	23.3%;	Score 532.2;	DB 12;	Length 561;
Best Local Similarity	98.6%;	Pred. No. 3e-89;		
Matches 545;	Conservative 0;	Mismatches 7;	Indels 1;	Gaps 1;
QY	1651	AGGCCTTGTTACATTGGCAATACAGTCCAAATGTGAATGATGATGCTCTGGTTATGT	1710	
Db	561	AGGCCTTGTTACATTGGCAATACAGTCCAAATGTGAATGATGATGCTCTGGTTATGT	502	
QY	1711	GGGGAGTCTGATGCGAGGCTTTTGAGTTCGTAGGGGTGTCCTAGTCTTCAGAACT	1770	
Db	501	GGGGAGTCTGATGCGAGGCTTTTGAGTTCGTAGGGGTGTCCTAGTCTTCAGAACT	442	
QY	1771	TGAATGAGAGGGTGTATATTTTCAGTGAACGTGCACTTGTCTGGCTGCAACACAAT	1830	
Db	441	TGAATGAGAGGGTGTATATTTTCAGTGAACGTGCACTTGTCTGGCTGCAACACAAT	382	
QY	1831	GACTTCTCTAGTACTTGTGGGTGCAAGGTTATGTTGATATCTCCATCTGGAAGTATCT	1890	
Db	381	GACTTCTCTAGTACTTGTGGGTGCAAGGTTATGTTGATATCTCCATCTGGAAGTATCT	322	
QY	1891	TTTGGTAATGGCTCGAACCCCTTTTGGAAACATTTAGTTGATTCCTCTAGAAAGTGGCTAC	1950	
Db	321	TTTGGTAATGGCTCGAACCCCTTTTGGAAACATTTAGTTGATTCCTCTAGAAAGTGGCTAC	262	

QY	1951	GAATACCAATCCAGATCAGACTGTAGTTGTTGAGCATCTCTGCTCATATTTCTTGCATATTA	2010		
Db	261	GAATACCAATCCAGATCAGACTGTAGTTGTTGAGCATCTCTGCTCATATTTCTTGCATATTA	202		
QY	2011	TTCTCTTTCAGGGCCAGAGATCAGATTTTCCAGATACCTGTTGCTTGGACACTGCCAC	2070		
Db	201	TTCTCTTTCAGGGCCAGAGATCAGATTTTCCAGATACCTGTTGCTTGGACACTGCCAC	142		
QY	2071	ATGCGTTTCATACCTAGAGGCCAGAGCTGTGTATATATACC-AGTTTTCTTTTGTCTTCT	2129		
Db	141	ATGCGTTTCATACCTAGAGGCCAGAGCTGTGTATATATACNNNTTTTCTTTTGTCTTCT	82		
QY	2130	TCTCCCTTTTCATATCTGTTTCTATGTTCTCTCTATTTGTAGTTTCATTTAGACAAT	2189		
Db	81	TCTCCCTTTTCATATCTGTTTCTATGTTCTCTCTATTTGTAGTTTCATTTAGACAAT	22		
QY	2190	TAGTCTTGTAAATA 2202			
Db	21	TAGTCTGTAAANA 9			
RESULT 10					
CF807291/c					
LOCUS					
DEFINITION					
pSHB022Xp18f USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation Phytophthora sojae cDNA clone SHB022P18 5, mRNA sequence.					
ACCESSION	CF807291	541 bp	mRNA	linear	EST 27-OCT-2003
VERSION	CF807291.1	GI:37995702			
KEYWORDS	EST.				
SOURCE	Phytophthora sojae				
ORGANISM	Phytophthora sojae				
REFERENCE	1 (bases 1 to 541)				
AUTHORS	Tyler, B.				
TITLE	Unpublished (2003)				
JOURNAL	Contact: Tyler B				
COMMENT	Tyler lab				
FEATURES					
source	1..541				
/organism="Phytophthora sojae"					
/mol_type="mRNA"					
/db_xref="taxon:67593"					
/clone="SHB022P18"					
/tissue_type="mycelium"					
/cell_line="P6497"					
/dev_stage="48 hr. post infection stage"					
/lab_host="Soybean plant"					
/clone_lib="USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation"					
/note="vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"					
ORIGIN					
Query Match	23.0%;	Score 526;	DB 14;	Length 541;	
Best Local Similarity	100.0%;	Pred. No. 4.3e-88;			
Matches 526;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1763	CAGAAACTTGAATGAGAGGGTGTATTTTTCAGTGAACGTGCACTTGTGGTGGTCA	1822		
Db	541	CAGAAACTTGAATGAGAGGGTGTATTTTTCAGTGAACGTGCACTTGTGGTGGTCA	482		

1823 ACACAAATGACTCTCTCTAGGTAAGTCTGGGTCGAAGGTTATGGTATCTCTCATCTGGA 1882  
Db |||||  
481 ACACAAATGACTCTCTCTAGGTAAGTCTGGGTCGAAGGTTATGGTATCTCTCATCTGGA 422  
Qy |||||  
1883 CQTGATCTTTTGGTAATGCTGACCCCTTTTGAACATTGAGTTGATTCCTTCTAGAAAG 1942  
Db |||||  
421 CQTGATCTTTTGGTAATGCTGACCCCTTTTGAACATTGAGTTGATTCCTTCTAGAAAG 362  
Qy |||||  
1943 GTGGCTACGAATACCAATACGATGAGACTGTAGTTGTTGAGCATCCTCTCATATCTTT 2002  
Db |||||  
361 GTGGCTACGAATACCAATACGATGAGACTGTAGTTGTTGAGCATCCTCTCATATCTTT 302  
Qy |||||  
2003 GCATATATCTTCTGACGGCAGAGATCAGATTTTCCAGATCTGTTGCTTGGAC 2062  
Db |||||  
301 GCATATATCTTCTGACGGCAGAGATCAGATTTTCCAGATCTGTTGCTTGGAC 242  
Qy |||||  
2063 ACTGCCACATCGCTTGATACCTAGAGCCAGAGCTGTGTATATATACCAAGTTTCTTTTG 2122  
Db |||||  
241 ACTGCCACATCGCTTGATACCTAGAGCCAGAGCTGTGTATATATACCAAGTTTCTTTTG 182  
Qy |||||  
2123 TTTTCTCTCTCCCTTTTATATGCTGTCTATGTTCTCTCTCTATTTGATGATTTT 2182  
Db |||||  
181 TTTTCTCTCTCCCTTTTATATGCTGTCTATGTTCTCTCTCTATTTGATGATTTT 122  
Qy |||||  
2183 AGCAATATGCTCTTAAATGAGCTGTTTCTTCAATTTGAAATCTGAAAGCTTCCCTTA 2242  
Db |||||  
121 AGCAATATGCTCTTAAATGAGCTGTTTCTTCAATTTGAAATCTGAAAGCTTCCCTTA 62  
Qy |||||  
2243 ACCTATTGGCTCCCTTAAATGAGCTGTTTCTTCAATTTGAAATCTGAAAGCTTCCCTTA 2288  
Db |||||  
61 ACCTATTGGCTCCCTTAAATGAGCTGTTTCTTCAATTTGAAATCTGAAAGCTTCCCTTA 16

RESULT 11  
BI974579  
LOCUS  
DEFINITION  
bai70d03.y1 Gm-cl068 Glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl068-3990 5' similar to TR:004197 004197 HYPOTHETICAL  
PROTEIN. i, mRNA sequence.  
ACCESSION  
BI974579.1 GI:16348984  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Glycine max (soybean)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 538)  
REFERENCE  
AUTHORS  
Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V.,  
Khanna, A., Bolla, B., Marras, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wiley, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 410.  
Location/Qualifiers  
1..538  
/organism="Glycine max"  
/mol\_type="mRNA"

/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl068-3990"  
/tissue\_type="Leaf, drought stressed, 1 month old plants,  
greenhouse grown"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl068"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI. The cDNA library was constructed from mRNA isolated  
from drought stressed leaf tissue of the cultivar Williams  
82. The month old greenhouse grown plants were deprived of  
water for 3 days prior to harvesting the stressed leaf  
tissue. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

ORIGIN  
Query Match 22.5%; Score 515.6; DB 12; Length 538;  
Best Local Similarity 97.4%; Pred. No. 3.8e-86;  
Matches 524; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
Qy 790 CAAGAGTTTAAAGAGTCTCTTTTGGAGGAAAGCTCAATCTTTGAGAGGACGAGAAATG 849  
Db 1 CAGGAGTTTAAAGAGTCTCTTTTGGAGGAAAGCTCAATCTTTGAGAGGACGAGAAATG 60  
Qy 850 GTTACACGAGCTTGTCTTGAATAATACAGTCTCTTGGAGCTCTCAATTTTACTTGACAG 909  
Db 61 GTTACACGAGCTTGTCTTGAATAATACAGTCTCTTGGAGCTCTCAATTTTACTTGACAG 120  
Qy 910 CATTGCTGTGTAAGATTGAGGACCTTGAACCTTTTAGCTTAAATTTGCCCAACTTAGT 969  
Db 121 TATTGCTGTGTAAGATTGAGGACCTTGAACCTTTTAGCTTAAATTTGCCCAACTTAGT 180  
Qy 970 GTCTGTGAACCTTACTGACTGTGAAATATCTGGAATCTTGTGAATCTCTTAAAGCATGCTC 1029  
Db 181 GTCTGTGAACCTTACTGACTGTGAAATATCTGGAATCTTGTGAATCTCTTAAAGCATGCTC 240  
Qy 1030 TGGCTGTGAAGAGTCTTGTGGAGGACCTTCAACAGAGGAAACCAAGAGATCTCTGCTAT 1089  
Db 241 TGGCTGTGAAGAGTCTTGTGGAGGACCTTCAACAGAGGAAACCAAGAGATCTCTGCTAT 300  
Qy 1090 ATCAATTACAGCAAGTATGTCGATTGGTGTAAACATATATTGGAAGAAATGAGTTGCC 1149  
Db 301 ATCAATTACAGCAAGTATGTCGATTGGTGTAAACATATATTGGAAGAAATGAGTTGCC 360  
Qy 1150 CATTGCTGTGTAAGATTGAGGACCTTGAACCTTTTAGCTTAAATTTGCCCAACTTAGT 1209  
Db 361 CATTGCTGTGTAAGATTGAGGACCTTGAACCTTTTAGCTTAAATTTGCCCAACTTAGT 420  
Qy 1210 CACGAGGAGTCAATGTATGTAATCAAGGTGTCGAATCTGGAAGTCTCTTGAGCAAG 1269  
Db 421 CACGAGGAGTCAATGTATGTAATCAAGGTGTCGAATCTGGAAGTCTCTTGAGCAAG 480  
Qy 1270 GAATGTAATTGGAGATAGAGGTTAGAGGTTCTTGGTCTGTTGTTGTAAGAGGCTTAAA 1327  
Db 481 GAATGTAATTGGAGATAGAGGTTAGAGGTTCTTGGTCTGTTGTTGTAAGAGGCTTAAA 538

RESULT 12  
CA820349  
LOCUS  
DEFINITION  
saub7g05.y1 Gm-cl048 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl048-3009 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.  
i, mRNA sequence.  
ACCESSION  
CA820349  
VERSION  
KEYWORDS  
SOURCE  
Glycine max (soybean)

ORGANISM	Glycine max
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
AUTHORS	1 (bases 1 to 550) Shoemaker R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cu@resgen.com www.resgen.com Seq primer: -4ORP from Gibco High quality sequence stop: 426.
FEATURES	Location/Qualifiers 1..550 /organism="Glycine max" /mol_type="mRNA" /db_xref="taxon:3847" /clone="SOYBEAN CLONE ID: Gm-cl048-3009" /tissue_type="whole seedlings of greenhouse grown plants" /dev_stage="1 week old" /lab_host="DH10B" /clone_lib="Gm-cl048" /notes="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; The Clark NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 1 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). The library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."
source	
ORIGIN	Query Match 21.6%; Score 493.4; DB 14; Length 550; Best Local Similarity 94.0%; Pred. No. 5.3e-82; Matches 529; Conservative 0; Mismatches 21; Indels 13; Gaps 1; Qy 1592 GCGTGTGACAGCTGAGAGATTTCCTATATCTCAGGGTGGCGGGTGACGTGATGTA 1651 Db 1 GCGTGTAAACAGCTGAGAGATTTCCTATATCTCAGGGTGGCGGGTGACCGATGTA 60 Qy 1652 GGCCTTGTTACATTGGACATACAGTCCAAATGTGAGATGCTGCTGGTTATGTG 1711 Db 61 GGTCTTGGTTACATTGGACATACAGTCCAAATGTGAGATGCTGCTGGTTATGTG 120 Qy 1712 GGGGAGTCTGATGACGGCTTTTGGAGTTTCGCTTAAGGGTGTCTCTTCAGAAACTT 1771 Db 121 GGGGAGTCTGATGACGGCTTTTGGAAATTCCTAAAGGGTGTCTCTTCAGAAACTA 180 Qy 1772 GAAATGAGAGGGTGTATTTTTCATGTAAGCTGCACCTGTCTGGTGGCAACAATTTG 1831 Db 181 GAAATGAGAGGGTGTCTATTTTTCAGTGAACGTGCACCTGTCTGGTGGCAACAATTTG 240

/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from whole seedlings of 2-3 week old greenhouse grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

ORIGIN		Query Match	21.4%;	Score 489;	DB 9;	Length 493;
		Best Local Similarity	100.0%;	Pred. No. 3.5e-81;		
		Matches 489;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1007	GTGAATCTTTAAGCATGCTCTGCCTGGAAGAGTTTGTGAGGCGACCTACACGAG	1066			
Db	5	GTGAATCTTTAAGCATGCTCTGCCTGGAAGAGTTTGTGAGGCGACCTACACGAG	64			
QY	1067	GAACAGAAAGATCTCTGTATATCATATACAGCAAGTTATGTGATGGGTTTAAACA	1126			
Db	65	GAACAGAAAGATCTCTGTATATCATATACAGCAAGTTATGTGATGGGTTTAAACA	124			
QY	1127	TATATTGAAAGATGAGTTGCCATTTGTTCATGTTTGCAGCGGTACTAAAAAATTG	1186			
Db	125	TATATTGAAAGATGAGTTGCCATTTGTTCATGTTTGCAGCGGTACTAAAAAATTG	184			
QY	1187	GATCTCTCTATGCAATGTAGACAGGAGTATCTTATCTTAATCAAGAGTGTCCA	1246			
Db	185	GATCTCTCTATGCAATGTAGACAGGAGTATCTTATCTTAATCAAGAGTGTCCA	244			
QY	1247	ATCTGGAAGTCTTGGACAGAGGATGTAATTTGAGATAGAGGTTAGAGTTCTTGGT	1306			
Db	245	ATCTGGAAGTCTTGGACAGAGGATGTAATTTGAGATAGAGGTTAGAGTTCTTGGT	304			
QY	1307	CGTTGTTGTAAGAGGCTTAAAGAGGCTTAGGATTAAGAGGCGGATGATCAAGGAATG	1366			
Db	305	CGTTGTTGTAAGAGGCTTAAAGAGGCTTAGGATTAAGAGGCGGATGATCAAGGAATG	364			
QY	1367	GAGGATGAAGAGGTTACTGTGTCCTATAGAGGCTTAATAGCCTTGTCAAGGCGTGTCA	1426			
Db	365	GAGGATGAAGAGGTTACTGTGTCCTATAGAGGCTTAATAGCCTTGTCAAGGCGTGTCA	424			
QY	1427	GAGCTTGAATACATGCTGCTTTATGTTCTGATATTAACAATGCATCTCTGGAACATATT	1486			
Db	425	GAGCTTGAATACATGCTGCTTTATGTTCTGATATTAACAATGCATCTCTGGAACATATT	484			
QY	1487	GGAACTCAC	1495			
Db	485	GGAACTCAC	493			

RESULT 14  
BQ630213  
LOCUS  
DEFINITION BQ630213.1 555 bp mRNA linear EST 21-OCT-2002  
Gm-c1045-3248 5' similar to TR:004197 004197 HYPOTHETICAL PROTEIN.  
i, mRNA sequence.  
ACCESSION BQ630213  
VERSION BQ630213.1 GI:21677862  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
REFERENCE 1 (bases 1 to 555)

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
McCann, R., Waterston, R., and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

TITLE  
JOURNAL  
COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: ccu@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 434.  
Location/Qualifiers

FEATURES

1..555  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1045-3248"  
/tissue\_type="Hypocotyl, 9-10 day old etiolated seedlings"  
/lab\_host="DH10B"  
/clone\_lib="Gm-c1045"  
/note="Vector: pBluescriptII SK+; Site 1: EcoRI; Site 2:  
XhoI; This cDNA library was constructed from mRNA isolated  
from etiolated hypocotyl tissue of 9-10 day old seedlings  
of the cultivar Williams 82. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) primer with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by digestion with EcoRI and XhoI. The cDNA  
fragments were directionally cloned into the EcoRI-XhoI  
restriction site of the pBluescript vector. The ligated  
cDNA fragments were transformed into DH10B host cells  
(Gibco, BRL). This library was constructed by Dr. Randy  
Shoemaker."

ORIGIN

		Query Match	20.8%;	Score 476.2;	DB 13;	Length 555;
		Best Local Similarity	93.3%;	Pred. No. 8.7e-79;		
		Matches 525;	Conservative 0;	Mismatches 23;	Indels 15;	Gaps 2;
QY	1623	ATCTCAGGCGTGGCGGTTGACTGATGTAGGCTTGGTTGATTCGACATTCAGTCCAA	1682			
Db	6	AGCTCAGGCGTGGCGGTTGACCGATGTAGGCTTGGTTTACATTCGACATTCGACAA	65			
QY	1683	ATGTGAGATGGATGCTGCTTGGTTATGTGGGGAGTCTGATCAGGCGCTTTGGAGTTGG	1742			
Db	66	ATGTGAGATGGATGCTGCTTGGTTATGTGGGGAGTCTGATCAGGCGCTTTGGAAATCT	125			
QY	1743	CTAAGGGGTGCTCCTAGTCTTCAGAACTTGAATGAGAGGGTGTATTTTCAGTGAAC	1802			
Db	126	CTAAGGGGTGCTCCTAGTCTTCAGAACTTGAATGAGAGGGTGTATTTTCAGTGAAC	185			
QY	1803	GTGCACTTCTGCTGGCTGCAACACATTCGATCTTCTTAGTACTTGTGGTGCAGGTT	1862			
Db	186	GTGCACTTCTGCTGGCTGCAACACATTCGATCTTCTTAGTACTTGTGGTGCAGGTT	245			
QY	1863	ATGGTGTATCTCCATCTGACGCTGATCTTTTGGTATGGCTCGACCTTTTGAACATTG	1922			
Db	246	ATGGTGTATCTCCATCTGACGCTGATCTTTTGGCAATGGCTCGCCCTTTTGAACATTG	305			
QY	1923	AGTTGATTCCTCTTAGAAAGGTTGGCTACGAATACCAATCCAGATGAGACTGATGTTG	1982			
Db	306	AGTTGATTCCTCTTAGAAAGGTTGGCTATGAATACCAATTCAGATGAGCGTATGTTG	365			



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 22:56:50 ; Search time 77 Seconds  
(without alignments)  
2223.686 Million cell updates/sec

Title: US-10-009-791-22  
Perfect score: 3180  
Sequence: 1 TKTSAFLFTLSLRNMTEE.....QRSDFPDVVPLDPTACVDT 606

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3180	100.0	606	3 AAB23458	Aab23458 Soybean C
2	2126.5	66.9	592	3 AAB23466	Aab23466 Arabidops
3	1854.5	58.3	597	3 AAB23457	Aab23457 Rice COI1
4	1245.5	39.2	429	3 AAB23456	Aab23456 Corn COI1
5	818	25.7	585	3 AAG30234	Aag30234 Arabidops
6	812.5	25.6	236	3 AAB23449	Aab23449 Rice COI1
7	812.5	25.6	236	3 AAB23449	Aap01920 O. sativa
8	716.5	22.3	221	3 AAB23461	Aab23461 Wheat COI
9	678	21.3	522	3 AAG30235	Aag30235 Arabidops
10	613.5	19.3	177	3 AAP01922	Aap01922 T. aestiv
11	613.5	19.3	177	3 AAB23451	Aab23451 Wheat COI
12	579	18.2	490	3 AAG30236	Aag30236 Arabidops
13	555	17.5	108	3 AAB23450	Aab23450 Soybean C
14	555	17.5	108	3 AAP01921	Aap01921 G. max CO
15	476	15.0	342	3 AAG36270	Aag36270 Arabidops
16	475.5	15.0	194	3 AAB23448	Aab23448 Corn COI1
17	460	14.5	186	3 AAB23455	Aab23455 Corn COI1
18	460	14.5	186	3 AAP01926	Aap01926 Z. mays C
19	303.5	9.5	242	3 AAG36272	Aag36272 Arabidops
20	256.5	8.1	610	3 AAG12663	Aag12663 Arabidops
21	250.5	7.9	159	3 AAB23459	Aab23459 Wheat COI
22	249.5	7.8	610	3 AAG52202	Aag52202 Arabidops
23	239	7.5	243	3 AAG36271	Aag36271 Arabidops
24	213	6.7	405	3 AAG43539	Aag43539 Arabidops
25	202	6.4	721	4 AAG78364	Aag78364 GRR1 homo

26	196	6.2	699	4 ABB71219	Abb71219 Drosophil
27	192	6.0	134	3 AAB23460	Aab23460 Wheat COI
28	187.5	5.9	628	4 AAG78365	Aag78365 Arabidops
29	179	5.6	422	5 ABJ10600	ABJ10600 Human nov
30	179	5.6	436	6 ABR43242	Abr43242 Human PMM
31	175	5.5	423	4 AAB92961	Aab92961 Human pro
32	173	5.4	423	4 AAB92791	Aab92791 Human pro
33	172.5	5.4	464	4 ABB64154	Abb64154 Drosophil
34	172	5.4	423	6 ABO07181	AbO07181 Human P53
35	172	5.4	483	3 AAB40624	Aab40624 Human ORF
36	172	5.4	483	3 AAY83087	Aay83087 F-box pro
37	172	5.4	483	4 AAB48305	Aab48305 Human ZF2
38	172	5.4	483	5 AAO22473	Aao22473 Human F-b
39	170.5	5.4	466	4 AAB48290	Aab48290 Human ZF1
40	164.5	5.2	456	2 AAY02274	Aay02274 A F-box p
41	164.5	5.2	456	4 AAE08046	Aae08046 Human ful
42	164.5	5.2	456	7 AAG39654	Aag39654 Human P-b
43	163.5	5.1	349	3 AAG43540	Aag43540 Arabidops
44	161.5	5.1	693	4 ABG70904	Abg70904 Arabidops
45	160.5	5.0	437	3 AAY83090	Aay83090 F-box pro

ALIGNMENTS

RESULT 1  
AAB23458  
ID AAB23458 standard; protein; 606 AA.  
XX  
AC AAB23458;  
XX  
DT 12-FEB-2001 (first entry)  
XX  
DE Soybean COI1 protein #2.  
XX  
KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
KW A. thaliana coronatine-induced; COI1; herbicides; EST;  
KW expressed sequence tag; sg94c.pk003.k23.fis.  
XX  
OS Glycine max.  
XX  
PN WO200068406-A2.  
XX  
PD 16-NOV-2000.  
XX  
PF 03-MAY-2000; 2000WO-US011956.  
XX  
PR 07-MAY-1999; 99US-0133041P.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX  
PI Cairni PG, Famodu OO, Lee J, Miao G, Maxwell CA;  
XX  
DR WPI; 2000-687649/67.  
XX  
PS N-PSDB; AAA95063.  
XX  
PT New nucleic acid sequences encoding new disease resistance factors,  
PT useful for producing plants with increased resistance to pathogens and to  
PT screen for herbicides.  
XX  
PS Claim 10; Fig 1; 74pp; English.  
XX  
CC Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
CC libraries were found to be similar to cDNA encoding the Arabidopsis  
CC thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence  
CC is protein encoded by cDNA from the soybean clone sg94c.pk003.k23.fis  
CC which is homologous to COI1. The COI1 and LLS1 DNAs of the invention may be used  
CC to alter the expression of COI1 and LLS1 protein in cells, particularly  
CC to produce transgenic plants with increased systemic resistance to a wide  
CC range of pathogens. COI1 and LLS1 proteins may be used to identify  
CC inhibitors of these proteins, which may be useful as herbicides  
XX



[illegible]



SQ Sequence 429 AA;  
 Query Match 39.2%; Score 1245.5; DB 3; Length 429;  
 Best Local Similarity 57.5%; Pred. No. 6.9e-123;  
 Matches 248; Conservative 69; Mismatches 95; Indels 19; Gaps 7

QY	176 KSLRVLFLESSILKSDG-EWHLALANNTVLETINFLTDIAVVKIEDIELLAKNCPL 234
DB	6 RGLTFLFLESTIDEKENDEWIRELATNSVLTFNFFLTDLR-ASPEYLTLVRNCQL 64
QY	235 VSVKLTCIDILLDNVFFKHSALEEFCCGTNYEEPE-----RYSAISLPKLCRLGLTYI 2899
DB	65 KTKISECFMPLDLSLFRTAQTLOBFAGSPFEQQGVASRNENYIFPPSLHRLSLLYM 124
QY	290 GKNELPVFMAAVLKLDLLYAMLDTEHDCHMLIORCENLEVLETNRNVIGDRGLEVLGR 3494
DB	125 GTNDMQLFPYATALKLDLQFTFLSTEDHCQIVQRCSNLETLEVRDVIGDRGLQVVAQT 184
QY	350 CKIKLRIRBERGDDQ-GMEDEBEGTVSHRGIIALSOGCSELEYMAVYVSDITNASLEHG 408
DB	185 CKKHLRLVRERGGDDQGLEDQBGRISQVGLMAIAQGCPELTYAHVHSYDTINAALBAVG 2444
QY	409 THLNKICDFRLVLLDHEEKITDLPDNGVRALLRCCKLRRFALYLRRCGLTDVGLGVIG 468
DB	245 TCSKNLNDPRLVDREAHITDPDNGVRALLRCCTKLRRPAFYVRFGALSDVGLGVYG 411
QY	469 QYSPNVRWMLLVGVGESDAGLLEFAFKGPCPSLOKLEMRCCLFFSERALAVAATOLTSLRYL 528
DB	305 EFKSIHYMLLGNVESDNGIQLSKGCPSLOKLEVRGCL-FSEHALAALAQLKSRLYL 363
QY	529 WVQGYGVSPGRDLLVMARPFNNIELIPSRKVATNTNP-ETVVVEHPAHILAYISLAGQ 589
DB	364 WVQGFSSPTGTDIMAWRPFWNIYIV-----PDQEPCEPHKRQILAYISLAGR 411
QY	588 RSDEPDPTVWL 598
DB	415 RTDCPPSVTL 425

RESULT 5

ARG30234

ID ARG30234 standard; protein; 585 AA.

AC AAG30234;

XX

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 36110.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

```
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 25.7%; Score 818; DB 3; Length 585;
Best Local Similarity 34.6%; Pred. No. 3.4e-77; Indels 32; Gaps 12;
Matches 200; Conservative 109; Mismatches 237;

Qy 32 VLDVPIYDDPKORDAVSOVCREWYELDSITRKHVTIALCYTTTPARLRFRPHLESK 91
Db 9 VLEHILSFIDSNEDNSVLSVCKSWPETERKTRKRVFVGNVCYAVPAVTRFFPEMRSLT 68
Qy 92 LKGPRAAMFNLIPEDEGWGHVTPWVKEISQVFDCKSLHFRMIVKSDLNLA---RDR 148
Db 69 LKGPFPADYNLVPDGGWGYAWPIEAMAAKSSLEIRMKRMVVTDECLEKIAASFKD- 127
Qy 149 GHVLHALKDKSCGFTTDLGFHICRCKSLRVLFLERSSILEKDGKWEHLALANNVLET 208
Db 128 ---FKVLVLTSCBGFSTIDGIAIAATCRNLKVLRECEIVEDLGDWLSYFPESSTLSV 184
Qy 209 LNFYLTDIADVVKIRDLELLAKNCNPNVSVKLTDCBILD-LYNFFKHASALBEFCGTYNE 267
Db 185 LDFSCLD-SEVKISDLRELVSRSFNLSKSLKLNPAVTLDGLVSLRLRCAPOLTELGTGFAA 243
Qy 268 E--PERYSAISLPAKLCR-----LGLTYIGKNELPIVFWFAVLKLLDLYAMLDTEHC 320
Db 244 QLKPEAFSKLSEAFSNCKQLQSLGLMDVLPFLPALYSVCPGLTSLNLSYATVRMDLV 303
Qy 321 MLIORCPNLEVTNRNVIQDRGLEVLGRCKRKLRLRIERGDDQDCQDEEGTVSHRGLI 380
Db 304 ELLRRCCKLQKLVNDLIEDKLEAVASYCKELRELVRFPSEPD--LDATNIPLEQGLV 361
Qy 381 ALSQGCSELEYMAVVSVDITNASLEHIGTHLKNLCDFRLVLD--HEKITDPLDNGVR 438
Db 362 FVSKGCRKLESVLYFCVQFTNAALFTIARKRPNLCKFELCVIEFPFAPDYKINEPLDKGFK 421
Qy 439 ALLRGCDKLRFPALVLRGGTLVDGLVIGQYSPNVRWMLLYGVGESDAGLEFAKGCP 498
Db 422 AIAEGCRDLRLSV---SGLUSDKAFKYGKHKVRLMSIAFAGDSLMHLHILSGCES 478
Qy 499 LQKLEMRCGLFFSERALAVAATQTLTSLRYLVQGYGVSPSGRDLVLMARFPWNILIPSR 558
Db 479 LKXLEIRDCP-FGDTALLLHAHAUKLETMSLWSSCFVSFGACKLLSOKMPLRNVEVI--- 534
Qy 559 KVATNTNPDETUVVEHPA-HILAYISLAGQSDPDPTV 595
Db 535 ----DEHPPESRPSPVPERIYIYRTVAGPRMDTPEFV 568

RESULT 6
AAB23449
ID AAB23449 standard; protein; 236 AA.
XX
AC AAB23449;
```

```

XX DT 12-FEB-2001 (first entry)
XX DE Rice CO11 protein.
XX KW Corn; rice; soybean; wheat; ATHCOR1, LUS1; disease resistance;
XX KW A. thaliana coronatine-induced; CO11; herbicides; EST;
XX KW expressed sequence tag; contig.
XX OS Oryza sativa.
XX FH Key Location/Qualifiers
FT Misc-difference 115 /label= Unknown
FT Misc-difference 219 /note= "Encoded by GG"
XX PN W0200068406-A2.
XX PD 16-NOV-2000.
XX PF 03-MAY-2000; 2000WO-US011956.
XX PR 07-MAY-1999; 99US-0133041P.
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PI Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;
XX DR WPI; 2000-687649/67.
XX DR N-PSDB; AAA95054.
XX PT New nucleic acid sequences encoding new disease resistance factors,
XX PT useful for producing plants with increased resistance to pathogens and to
XX PT screen for herbicides.
XX PS Claim 10; Page 44-45; 74pp; English.
XX CC Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA
XX CC libraries were found to be similar to cDNA encoding the Arabidopsis
XX CC thaliana CO11 protein and the Zea mays LUS1 protein. The present sequence
XX CC is protein encoded by cDNA from the rice contig which is homologous to
XX CC CO11. The CO11 and LUS1 proteins are involved in disease resistance. The
XX CC CO11 and LUS1 DNAs of the invention may be used to alter the expression
XX CC of CO11 and LUS1 protein in cells, particularly to produce transgenic
XX CC plants with increased systemic resistance to a wide range of pathogens.
XX CC CO11 and LUS1 proteins may be used to identify inhibitors of these
XX CC proteins, which may be useful as herbicides
XX SQ Sequence 236 AA;
    Query Match 25.6%; Score 812.5; DB 3; Length 236;
    Best Local Similarity 66.8%; Pred. No. 3e-77;
    Matches 157; Conservative 30; Mismatches 47; Indels 1; Gaps 1;
    QY 328 NLEVLTRNVIQGRGLVGRCKKRLRIRGDDDDQGMDEEGTVSHRGLIALSOGCS 387
    Db 1 NLLVLAVRNVIQGRGLGVVADTCKKLQRLVRVERGDDPGLEEQGGVSGVGLTTVAVGCR 60
    QY 388 ELEYAVVYSDITNASLEHIGTHLKNLCDPRLVLLDHEEKITDPLDNGVRALLRGCDKL 447
    Db 61 ELEYIAVYSDITNGALESIGTCKNLCDFRLVLLDREERITDPLDNGVRALLXGCTKL 120
    QY 448 RRFALYLRGGGLTVGLVIGQYSPNVNMLLYGVGESDAGLLEFAKGCPSLOKLEMRGC 507
    Db 121 RRFALYLRPGGLSDTGLVIGQYSGIIQYMLLNGVGETDGLTRFALGCENLRKLEIRSC 180
    QY 508 LFSERALAVAAQTLSRLVWQGVSPSGRDLVWAPFPWNIILIPSKVAT 562
    Db 181 C-FSEQALARAIRSMPSRLRYVWQGVKASKTGHDLMLMARFPWNIETPPRLVAT 234
    RESULT 7

```

```

AAP01920
ID AAP01920 standard; protein; 236 AA.
XX AC AAP01920;
XX DT 30-OCT-2001 (first entry)
XX DE O. sativa CO11.
XX KW CO11; rice; herbicide; disease-resistance factor.
XX OS Oryza sativa.
XX FH Key Location/Qualifiers
FT Misc-difference 115 /label= "unknown"
FT Misc-difference 219 /note= "encoded by ANG"
FT Misc-difference 219 /note= "encoded by GG"
XX PN W0200068406-A2.
XX PD 16-NOV-2000.
XX PF 03-MAY-2000; 2000WO-US011956.
XX PR 07-MAY-1999; 99US-0133041P.
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PI Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;
XX DR WPI; 2000-687649/67.
XX DR N-PSDB; AAN02372.
XX PT New nucleic acid sequences encoding new disease resistance factors,
XX PT useful for producing plants with increased resistance to pathogens and to
XX PT screen for herbicides.
XX PS Claim 10; Page 44-45; 74pp; English.
XX CC An isolated polynucleotide (I) comprising a sequence (Ia) encoding a
XX CC polypeptide (II) of at least 60 amino acids (aa) from any of 18 specified
XX CC sequences (II) in the specification, or the complements of (Ia), is
XX CC new. Herbicide. (I) encodes disease-resistance factor polypeptides (II)
XX CC and can be used, in sense or antisense orientation, to alter the
XX CC expression of (II) in cells, particularly to produce transgenic plants
XX CC with increased systemic resistance to a wide range of pathogens. (I) or
XX CC their fragments are also used for gene mapping (e.g. for plant breeding)
XX CC and to identify loss of function mutants. Polypeptides (II) encoded by
XX CC (I) are used to raise specific antibodies (for detection of (II)) and to
XX CC design and/or identify specific inhibitors of (II), potentially useful as
XX CC herbicides. This sequence is the rice CO11 protein created from a contig
XX CC of xlr2.pk0027.b4, rlon.pk039.p14 and rlon.pk0047.c5 as described in the
XX CC method of the invention
XX SQ Sequence 236 AA;
    Query Match 25.6%; Score 812.5; DB 3; Length 236;
    Best Local Similarity 66.8%; Pred. No. 3e-77;
    Matches 157; Conservative 30; Mismatches 47; Indels 1; Gaps 1;
    QY 328 NLEVLTRNVIQGRGLVGRCKKRLRIRGDDDDQGMDEEGTVSHRGLIALSOGCS 387
    Db 1 NLLVLAVRNVIQGRGLGVVADTCKKLQRLVRVERGDDPGLEEQGGVSGVGLTTVAVGCR 60
    QY 388 ELEYAVVYSDITNASLEHIGTHLKNLCDPRLVLLDHEEKITDPLDNGVRALLRGCDKL 447
    Db 61 ELEYIAVYSDITNGALESIGTCKNLCDFRLVLLDREERITDPLDNGVRALLXGCTKL 120
    QY 448 RRFALYLRGGGLTVGLVIGQYSPNVNMLLYGVGESDAGLLEFAKGCPSLOKLEMRGC 507
    Db 121 RRFALYLRPGGLSDTGLVIGQYSGIIQYMLLNGVGETDGLTRFALGCENLRKLEIRSC 180

```

QY 508 LFFSERALAAVATOLTSRLRYLWQGVSPSGRDLMLVWAFWFFNIELPSKXVAT 562  
Db 181 C-FSEQLARAIRSMPSRLRYVWQGVKASKTGHDMLMLWAFWFFNIELPFRRLVT 234  
RESULT 8  
AAB23461  
ID AAB23461 standard; protein; 221 AA.  
XX AC AAB23461;  
XX DT 12-FEB-2001 (first entry)  
XX DE Wheat CO11 protein #4.  
XX KW Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance;  
XX KW A. thaliana coronatine-induced; CO11; herbicides; EST;  
XX KW expressed sequence tag; wrln.pk0122.d3.fis.  
XX OS Triticum aestivum.  
XX PN WO200068406-A2.  
XX PD 16-NOV-2000.  
XX PF 03-MAY-2000; 2000WO-US011956.  
XX PR 07-MAY-1999; 99US-0133041P.  
XX PA (DUPO) DU PONT DE NEMOURS & CO E. I.  
XX PI Caiami PG, Famodu OO, Lee J, Miao G, Maxwell CA;  
XX WPI; 2000-687649/67.  
XX DR N-PSDB; AAA95066.  
XX PT New nucleic acid sequences encoding new disease resistance factors,  
XX PT useful for producing plants with increased resistance to pathogens and to  
XX PT screen for herbicides.  
XX PS Claim 10; Page 66-67; 74pp; English.  
XX CC Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
XX CC libraries were found to be similar to cDNA encoding the Arabidopsis  
XX CC thaliana CO11 protein and the Zea mays LLS1 protein. The present sequence  
XX CC is protein encoded by cDNA from the wheat clone wrln.pk0122.d3.fis which  
XX CC is homologous to CO11. The CO11 and LLS1 proteins are involved in disease  
XX CC resistance. The CO11 and LLS1 DNAs of the invention may be used to alter  
XX CC the expression of CO11 and LLS1 protein in cells, particularly to produce  
XX CC transgenic plants with increased systemic resistance to a wide range of  
XX CC pathogens. CO11 and LLS1 proteins may be used to identify inhibitors of  
XX CC these proteins, which may be useful as herbicides  
XX SQ Sequence 221 AA;  
Query Match 22.5%; Score 716.5; DB 3; Length 221;  
Best Local Similarity 66.1%; Pred. No. 4.3e-67;  
Matches 146; Conservative 22; Mismatches 44; Indels 9; Gaps 3;  
QY 378 GLIALSQGSELEYMAVYVDITNASLEHIGTHKNLCDFRVLDDHEEKITDPLDNGV 437  
Db 4 GLMAVAGCPDLEYMAVHVSDDITNALEAIGASKNLDFRVLDDHEEVIHELPLDNGV 63  
QY 438 RALLRGCDKLRREALYLRGLTVDGLGIVGYQSPNVRWMLLGVGSGDAGLLEFAKGP 497  
Db 64 RALLRGCTKLRRFAFYVRPCALSDIGLSYVGFEKTVRYMLNAGSGDDGLLAFARGCP 123  
QY 498 SLQKLEMRGCLFFSERALAAVATOLTSRLRYLWQGVSPSGRDLMLVWAFWFFNIELIPS 557  
Db 124 SLQKLELRSCC-FSERALAAVATOLTSRLRYLWQGVKASKTGHDMLMLWAFWFFNIELFIA- 181  
QY 558 RKVATNTNPDETVEVHPAILAYSLAGORSDFPDVTVP 598

Db 182 -----FNQDEP-CPEGQAQILAYSLAGARTDQPSVIPL 215  
RESULT 9  
AAG30235  
ID AAG30235 standard; protein; 522 AA.  
XX AC AAG30235;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 36111.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
XX PR 05-MAR-1999; 99US-0123180P.  
XX PR 09-MAR-1999; 99US-0123548P.  
XX PR 23-MAR-1999; 99US-0125788P.  
XX PR 29-MAR-1999; 99US-0126264P.  
XX PR 06-APR-1999; 99US-0126785P.  
XX PR 01-APR-1999; 99US-0127462P.  
XX PR 08-APR-1999; 99US-0128234P.  
XX PR 16-APR-1999; 99US-0128714P.  
XX PR 19-APR-1999; 99US-0129845P.  
XX PR 21-APR-1999; 99US-0130077P.  
XX PR 23-APR-1999; 99US-0130449P.  
XX PR 28-APR-1999; 99US-0130510P.  
XX PR 30-APR-1999; 99US-0130891P.  
XX PR 04-MAY-1999; 99US-0131445P.  
XX PR 05-MAY-1999; 99US-0132048P.  
XX PR 06-MAY-1999; 99US-0132407P.  
XX PR 07-MAY-1999; 99US-0132484P.  
XX PR 11-MAY-1999; 99US-0132485P.  
XX PR 14-MAY-1999; 99US-0132486P.  
XX PR 14-MAY-1999; 99US-0132487P.  
XX PR 14-MAY-1999; 99US-0132863P.  
XX PR 14-MAY-1999; 99US-0134218P.  
XX PR 14-MAY-1999; 99US-0134219P.  
XX PR 14-MAY-1999; 99US-0134221P.  
XX PR 14-MAY-1999; 99US-0134370P.  
XX PR 18-MAY-1999; 99US-0134768P.  
XX PR 19-MAY-1999; 99US-0134941P.  
XX PR 20-MAY-1999; 99US-0135124P.  
XX PR 21-MAY-1999; 99US-0135353P.  
XX PR 24-MAY-1999; 99US-0135629P.  
XX PR 25-MAY-1999; 99US-0136021P.  
XX PR 27-MAY-1999; 99US-0136392P.  
XX PR 01-JUN-1999; 99US-0136782P.  
XX PR 03-JUN-1999; 99US-0137222P.  
XX PR 04-JUN-1999; 99US-0137528P.  
XX PR 07-JUN-1999; 99US-0137502P.  
XX PR 08-JUN-1999; 99US-0137724P.  
XX PR 10-JUN-1999; 99US-0138094P.  
XX PR 10-JUN-1999; 99US-0138540P.  
XX PR 14-JUN-1999; 99US-0138847P.  
XX PR 16-JUN-1999; 99US-0139119P.  
XX PR 16-JUN-1999; 99US-0139452P.  
XX PR 17-JUN-1999; 99US-0139453P.  
XX PR 18-JUN-1999; 99US-0139492P.  
XX PR 18-JUN-1999; 99US-0139454P.



PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	21-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	27-AUG-1999;	99US-0151080P.
PR	23-JUN-1999;	99US-0140353P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140354P.	PR	31-AUG-1999;	99US-0151438P.
PR	24-JUN-1999;	99US-0140695P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140823P.	PR	07-SEP-1999;	99US-0152363P.
PR	29-JUN-1999;	99US-0140991P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0141287P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141842P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0142154P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142055P.	PR	20-SEP-1999;	99US-0154039P.
PR	06-JUL-1999;	99US-0142390P.	PR	22-SEP-1999;	99US-0155139P.
PR	08-JUL-1999;	99US-0142803P.	PR	23-SEP-1999;	99US-0155486P.
PR	09-JUL-1999;	99US-0142920P.	PR	24-SEP-1999;	99US-0155659P.
PR	12-JUL-1999;	99US-0142977P.	PR	28-SEP-1999;	99US-0156458P.
PR	13-JUL-1999;	99US-0143342P.	PR	29-SEP-1999;	99US-0156596P.
PR	14-JUL-1999;	99US-0143624P.	PR	04-OCT-1999;	99US-0157117P.
PR	15-JUL-1999;	99US-0144005P.	PR	05-OCT-1999;	99US-0157533P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0157865P.
PR	19-JUL-1999;	99US-0144325P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144331P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144332P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159337P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160980P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160981P.
PR	27-JUL-1999;	99US-0145913P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161405P.
PR	28-JUL-1999;	99US-0145951P.	PR	25-OCT-1999;	99US-0161406P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161361P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161920P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161922P.
PR	05-AUG-1999;	99US-0147192P.	PR	28-OCT-1999;	99US-0161993P.
PR	05-AUG-1999;	99US-0147260P.	PR	29-OCT-1999;	99US-0162142P.
PR	06-AUG-1999;	99US-0147303P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			

Query Match

21.3%; Score 678; DB 3; Length 522;

Best Local Similarity 33.7%; Pred. No. 2.2e-62;

Matches 176; Conservative 96; Mismatches 219; Indels 32; Gaps 12;

Qy	87	LESLKLGKPRAAFMNLPEDWGHVTPWVKELISQYFDCIKSLHFRMIVKDSLOJLA-	145
Db	1	MRSUTLKGKPHFADYNLPDGGWGYAMPWIEAMAKSSLEETRMKRMVVTDECLEKIAA	60
Qy	146	--RDRGHVLHAKDKCSTGTTDGLFHIGRCKSLRVLFLEESSILEKDGELHLELNN	203
Db	61	SFKD----FKVLVLTSCGEGFTDGIATAATCRNLRLVLELRECIVEDLGGDWLSYFFESS	116

QY	204	TVLETLNFIYLTDAVVKIEDLELLAKNCPNLVSUKLTDCEILD-LVNPFKHASALEEPCG	262
DB	117	TSLVSLDFSCILD-SEVKISDLERLVSRSPNLKSLKLNPAVTLTDLGLVSLRCAPQLTELTG	175
QY	263	GTYNEE--PERYSAISLPKLCR-----LGLTYIGKNELPIVVFVFAAIVLKKLIDLIYAML	315
DB	176	GSPAQLKPEAFSKLSEAFSNCKQLQSLSGLMDVLPEVLPALYSVCPGLTSLNLSYATVR	235
QY	316	TEDHMLIQRCPNLVLETRNVIGDRGLEVLGRCCRLKRLRIERGGDDDDQGWEDDEGIVS	375
DB	236	MPDLVELLRCSKQLKWMVMDLIEDKLEAVASYCKELRELRFVFPSEPD--LDATNIPLT	293
QY	376	HRGLIALSQSCSELEYMAVYVSDITNASLEHIGTHLKNLDCDFRLVLLD--HEEKITDPL	433
DB	294	EQGLVFSVSGCRKLESVLYFCQFTNALFTIARKPNLKCFLCVIEFFAPDYKTNEL	353
QY	434	DNGVALLRGCDKLRREFALYLRGGITDVLGYIGQYSPNVRWMLLGYVGSDDAGLLBFA	493
DB	354	DKGFKAIABGCDRLRLSV--SGLLSDKAFYIGKHAKVEMLSIAFAGSDMLHLHL	410
QY	494	KGCPSLQKLEMRGCLFFSERALAVAATQTSRYLWVQYGVSPSGRDLVWVAPFWNIE	553
DB	411	SGCESLKKLEIRDCP-FGDTALLEHAAKLETRSLWMSGCFVSFGACKLLSQMPRLNVE	469
QY	554	LIPSRKVATNTPDETIVVVEHPA-HILAYYSLAGORSDFPDIV	595
DB	470	VI-----DEHPPESSPESPVERIYIYRTVAGPRMDTPEFV	505
RESULT 10			
ID	AAP01922	AAP01922 standard; protein; 177 AA.	
XX	AC	AAP01922;	
XX	DT	30-OCT-2001 (first entry)	
XX	DE	T. aestivum COIL.	
XX	KW	COIL; wheat; herbicide; disease-resistance factor.	
XX	OS	Triticum aestivum.	
XX	XX	Key Location/Qualifiers	
FT	FT	Misc-difference 84 /note= "encoded by GATA"	
XX	PN	WO200068406-A2.	
XX	PD	16-NOV-2000.	
XX	PF	03-MAY-2000; 2000WO-US011956.	
XX	PR	07-MAY-1999; 99US-0133041P.	
XX	PA	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX	PI	Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;	
XX	DR	WPI; 2000-687649/67.	
XX	DR	N-ESDB; AAN02374.	
XX	XX	New nucleic acid sequences encoding new disease resistance factors, useful for producing plants with increased resistance to pathogens and to screen for herbicides.	
XX	XX	Claim 10; Page 48-49; 74pp; English.	
XX	XX	An isolated polynucleotide (I) comprising a sequence (Ia) encoding a polypeptide (II) of at least 60 amino acids (aa) from any of 18 specified sequences, given in the specification, or the complements of (Ia), is new. Herbicide. (I) encodes disease-resistance factor polypeptides (II) and can be used, in sense or antisense orientation, to alter the	
CC	CC	expression of (II) in cells, particularly to produce transgenic plants with increased systemic resistance to a wide range of pathogens. (I) or their fragments are also used for gene mapping (e.g. for plant breeding) and to identify loss of function mutants. Polypeptides (II) encoded by (I) are used to raise specific antibodies (for detection of (II)) and to design and/or identify specific inhibitors of (II), potentially useful as herbicides. This sequence is the wheat COIL clone created from contigs wraln.pk0122.d3 and wlin.pk0018.f8 as described in the method of the invention	
XX	XX	SQ Sequence 177 AA;	
Query Match 19.3%; Score 613.5; DB 3; Length 177;			
Best Local Similarity 68.5%; Pred. No. 2.7e-56;			
Matches 122; Conservative 19; Mismatches 36; Indels 1; Gaps 1;			
QY	378	GLIALSQSCSELEYMAVYVSDITNASLEHIGTHLKNLDCDFRLVLLDHEEKITDPLDNGV	437
DB	1	GLMAVAGCPDLEYWAVHVSDITNALEAIGAFSKNLNDFRLVLLDREVHITELDNGV	60
QY	438	RALLRGCDKLRREFALYLRGGITDVLGYIGQYSPNVRWMLLGYVGSDDAGLLBFAKGP	497
DB	61	RALLRGCTKLRRFAPYVRPGALSDLAFLXGFEFSKTVYVMLGNAGSGDDGLLAFARXCP	120
QY	498	SLQKLEMRGCLFFSERALAVAATQTSRYLWVQYGVSPSGRDLVWVAPFWNIELI	555
DB	121	SLQKLELRSCC-FSERALAAALQLKSURLWVQYKASPTGTDLMAMVRFWNIEFI	177
RESULT 11			
ID	AAB23451	AAB23451 standard; protein; 177 AA.	
XX	AC	AAB23451;	
XX	DT	12-FEB-2001 (first entry)	
XX	DE	Wheat COIL protein.	
XX	KW	Corn; rice; soybean; wheat; ATHCOR1; LLS1; disease resistance; A. thaliana coronatine-induced; COIL; herbicides; EST;	
XX	OS	Triticum aestivum.	
XX	XX	Key Location/Qualifiers	
FT	FT	Misc-difference 85 /note= "Encoded by ATTG"	
FT	FT	Misc-difference 89 /label= Unknown	
FT	FT	Misc-difference 118 /label= Unknown	
XX	PN	WO200068406-A2.	
XX	PD	16-NOV-2000.	
XX	PF	03-MAY-2000; 2000WO-US011956.	
XX	PR	07-MAY-1999; 99US-0133041P.	
XX	PA	(DUPO ) DU PONT DE NEMOURS & CO E I.	
XX	PI	Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;	
XX	DR	WPI; 2000-687649/67.	
XX	DR	N-PSDB; AAA95056.	
XX	XX	New nucleic acid sequences encoding new disease resistance factors, useful for producing plants with increased resistance to pathogens and to screen for herbicides.	
XX	XX	Claim 10; Page 47-48; 74pp; English.	

XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA  
CC libraries were found to be similar to cDNA encoding the Arabidopsis  
CC thaliana COII protein and the Zea mays LLS1 protein. The present sequence  
CC is protein encoded by cDNA from the wheat contig which is homologous to  
CC COII. The COII and LLS1 proteins are involved in disease resistance. The  
CC COII and LLS1 DNAs of the invention may be used to alter the expression  
CC of COII and LLS1 protein in cells, particularly to produce transgenic  
CC plants with increased systemic resistance to a wide range of pathogens.  
CC COII and LLS1 proteins may be used to identify inhibitors of these  
CC proteins, which may be useful as herbicides  
XX  
SQ Sequence 177 AA;

Query Match 19.3%; Score 613.5; DB 3; Length 177;  
Best Local Similarity 68.5%; Pred No. 2.7e-56;  
Matches 122; Conservative 19; Mismatches 36; Indels 1; Gaps 1;

QY 378 GLIALSQGSELEYMAVYVSDITNASLEHIGTHLKNLDCDFRLVLLDHEEKITDPLDNGV 437  
DB 1 GLMAVAGCPDLEYNAHVSDITNALEAIGAFSKNLNDFRLVLLDREVHITELPLDNGV 60  
QY 438 RALLGCDKLRFPALYLRGGTVDVLGVIQYSPNVRMLLYGVGSDAGLLEFAKQCP 497  
DB 61 RALLRGCTKLRFAFYVRPGALSDFLAFIXIGBFSKTVRYMLGNAGGSDGDLAFARXCP 120  
QY 498 SLQKLEMRGCLFESERALVAATOLTSRLYLWQYGVSPSGRDLVMARPPWNIELI 555  
DB 121 SLQKLEMRGCLFESERALVAATOLTSRLYLWQYGVSPSGRDLVMARPPWNIEFI 177

RESULT 12  
AAG30236  
ID AAG30236 standard; protein; 490 AA.  
XX AC AAG30236;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 36112.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
XX PR 05-MAR-1999; 99US-0123180P.  
XX PR 23-MAR-1999; 99US-0123548P.  
XX PR 23-MAR-1999; 99US-0125788P.  
XX PR 23-MAR-1999; 99US-0126264P.  
XX PR 23-MAR-1999; 99US-0126785P.  
XX PR 01-APR-1999; 99US-0127462P.  
XX PR 08-APR-1999; 99US-0128234P.  
XX PR 16-APR-1999; 99US-0128714P.  
XX PR 19-APR-1999; 99US-0129845P.  
XX PR 21-APR-1999; 99US-0130077P.  
XX PR 23-APR-1999; 99US-0130449P.  
XX PR 23-APR-1999; 99US-0130510P.  
XX PR 28-APR-1999; 99US-0130891P.  
XX PR 30-APR-1999; 99US-0131449P.  
XX PR 30-APR-1999; 99US-0132048P.  
XX PR 04-MAY-1999; 99US-0132407P.  
XX PR 05-MAY-1999; 99US-0132484P.  
XX PR 06-MAY-1999; 99US-0132485P.  
XX PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.

PR	22-JUN-1999;	99US-0145192P;
PR	23-JUN-1999;	99US-0145145P;
PR	23-JUN-1999;	99US-0145218P;
PR	23-JUN-1999;	99US-0145224P;
PR	26-JUN-1999;	99US-0145226P;
PR	27-JUN-1999;	99US-0145913P;
PR	27-JUN-1999;	99US-0145918P;
PR	27-JUN-1999;	99US-0145919P;
PR	28-JUN-1999;	99US-0145951P;
PR	02-AUG-1999;	99US-0146386P;
PR	02-AUG-1999;	99US-0146388P;
PR	02-AUG-1999;	99US-0146389P;
PR	03-AUG-1999;	99US-0147038P;
PR	04-AUG-1999;	99US-0147204P;
PR	04-AUG-1999;	99US-0147302P;
PR	05-AUG-1999;	99US-0147192P;
PR	05-AUG-1999;	99US-0147260P;
PR	06-AUG-1999;	99US-0147303P;
PR	06-AUG-1999;	99US-0147416P;
PR	09-AUG-1999;	99US-0147493P;
PR	09-AUG-1999;	99US-01474935P;
PR	09-AUG-1999;	99US-0148171P;
PR	11-AUG-1999;	99US-0148319P;
PR	12-AUG-1999;	99US-0148341P;
PR	13-AUG-1999;	99US-0148565P;
PR	13-AUG-1999;	99US-0148684P;
PR	16-AUG-1999;	99US-0149368P;
PR	17-AUG-1999;	99US-0149175P;
PR	18-AUG-1999;	99US-0149426P;
PR	20-AUG-1999;	99US-0149722P;
PR	20-AUG-1999;	99US-0149723P;
PR	20-AUG-1999;	99US-0149929P;
PR	23-AUG-1999;	99US-0149902P;
PR	23-AUG-1999;	99US-0149930P;
PR	25-AUG-1999;	99US-0150556P;
PR	26-AUG-1999;	99US-0150884P;
PR	27-AUG-1999;	99US-0151085P;
PR	27-AUG-1999;	99US-0151066P;
PR	27-AUG-1999;	99US-0151080P;
PR	31-AUG-1999;	99US-0151303P;
PR	31-AUG-1999;	99US-0151438P;
PR	01-SEP-1999;	99US-0151930P;
PR	07-SEP-1999;	99US-0152363P;
PR	07-SEP-1999;	99US-0153070P;
PR	13-SEP-1999;	99US-0153758P;
PR	15-SEP-1999;	99US-0154018P;
PR	16-SEP-1999;	99US-0154039P;
PR	20-SEP-1999;	99US-0154779P;
PR	22-SEP-1999;	99US-0155139P;
PR	23-SEP-1999;	99US-0155486P;
PR	24-SEP-1999;	99US-0155659P;
PR	28-SEP-1999;	99US-0156458P;
PR	29-SEP-1999;	99US-0156596P;
PR	04-OCT-1999;	99US-0157117P;
PR	05-OCT-1999;	99US-0157753P;
PR	08-OCT-1999;	99US-0158029P;
PR	08-OCT-1999;	99US-0158232P;
PR	12-OCT-1999;	99US-0158369P;
PR	13-OCT-1999;	99US-0158293P;
PR	13-OCT-1999;	99US-0159294P;
PR	13-OCT-1999;	99US-0159295P;
PR	14-OCT-1999;	99US-0159329P;
PR	14-OCT-1999;	99US-0159330P;
PR	14-OCT-1999;	99US-0159331P;
PR	14-OCT-1999;	99US-0159637P;
PR	18-OCT-1999;	99US-0159638P;
PR	18-OCT-1999;	99US-0159584P;
PR	21-OCT-1999;	99US-0160741P;
PR	21-OCT-1999;	99US-0160767P;
PR	21-OCT-1999;	99US-0160768P;
PR	21-OCT-1999;	99US-0160770P;
PR	21-OCT-1999;	99US-0160814P;
PR	21-OCT-1999;	99US-0160841P;

```

FT Misc-difference 97 /label= Unknown
FT /note= "Encoded by TTGC"
PN
XX WO200068406-A2.
XX
XX 16-NOV-2000.
XX
XX 03-MAY-2000; 2000WO-US011956.
XX
XX 07-MAY-1999; 99US-0133041P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;
XX
XX WPI; 2000-687649/67.
XX
XX N-PSDB; AAA95055.
XX
XX New nucleic acid sequences encoding new disease resistance factors,
XX useful for producing plants with increased resistance to pathogens and to
XX screen for herbicides.
XX
XX Claim 10; Page 45-46; 74pp; English.
XX
XX Expressed sequence tags (ESTs) from rice, wheat, soybean and corn cDNA
XX libraries were found to be similar to cDNA encoding the Arabidopsis
XX thaliana COI1 protein and the Zea mays LLS1 protein. The present sequence
XX is protein encoded by cDNA from the soybean clone sgs4c.pk003.k23 which
XX is homologous to COI1. The COI1 and LLS1 proteins are involved in disease
XX resistance. The COI1 and LLS1 DNAs of the invention may be used to alter
XX the expression of COI1 and LLS1 protein in cells, particularly to produce
XX transgenic plants with increased systemic resistance to a wide range of
XX pathogens. COI1 and LLS1 proteins may be used to identify inhibitors of
XX these proteins, which may be useful as herbicides
XX
XX Sequence 108 AA;

Query Match 17.5%; Score 555; DB 3; Length 108;
Best Local Similarity 96.3%; Pred. No. 2.1e-50;
Matches 104; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 29 VDVLDCVPIYDDPKDRDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRPPHLE 88
Db 1 VDVLDCVPIYDDPKDRDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRPPHLE 60

QY 89 SLKLKGPRAAMFNLIPEWGGHVTTPWVKEISQVFDCLKSLHFRMIV 136
Db 61 SLKLKGPRAAMFNLIPEWGGHVTTPWVKEISQVLRXKLSLHFRMIV 108

RESULT 14
AAP01921
ID AAP01921 standard; protein; 108 AA.
XX
XX AAP01921;
XX
XX 30-OCT-2001 (first entry)
XX
XX G. max COI1.
XX
XX COI1; soybean; herbicide; disease-resistance factor.
XX
XX Glycine max.
XX
XX Key Location/Qualifiers
XX Misc-difference 97 /label= unknown
XX /note= "encoded by TTGC"
XX
XX WO200068406-A2.
XX
XX 16-NOV-2000.
XX
XX 03-MAY-2000; 2000WO-US011956.
XX
XX 07-MAY-1999; 99US-0133041P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Caimi PG, Famodu OO, Lee J, Miao G, Maxwell CA;
XX
XX WPI; 2000-687649/67.
XX
XX N-PSDB; AAN02373.
XX
XX New nucleic acid sequences encoding new disease resistance factors,
XX useful for producing plants with increased resistance to pathogens and to
XX screen for herbicides.
XX
XX Claim 10; Page 45-46; 74pp; English.
XX
XX An isolated polynucleotide (I) comprising a sequence (Ia) encoding a
XX polypeptide (II) of at least 60 amino acids (aa) from any of 18 specified
XX sequences, given in the specification, or the complements of (Ia), is
XX new. Herbicide. (I) encodes disease-resistance factor polypeptides (II)
XX and can be used, in sense or antisense orientation, to alter the
XX expression of (II) in cells, particularly to produce transgenic plants
XX with increased systemic resistance to a wide range of pathogens. (I) or
XX their fragments are also used for gene mapping (e.g. for plant breeding)
XX and to identify loss of function mutants. Polypeptides (II) encoded by
XX (I) are used to raise specific antibodies (for detection of (II)) and to
XX design and/or identify specific inhibitors of (II), potentially useful as
XX herbicides. This sequence is the soybean COI1 clone sgs4c.pk003.k23 as
XX described in the method of the invention
XX
XX Sequence 108 AA;

Query Match 17.5%; Score 555; DB 3; Length 108;
Best Local Similarity 96.3%; Pred. No. 2.1e-50;
Matches 104; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 29 VDVLDCVPIYDDPKDRDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRPPHLE 88
Db 1 VDVLDCVPIYDDPKDRDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRPPHLE 60

QY 89 SLKLKGPRAAMFNLIPEWGGHVTTPWVKEISQVFDCLKSLHFRMIV 136
Db 61 SLKLKGPRAAMFNLIPEWGGHVTTPWVKEISQVLRXKLSLHFRMIV 108

RESULT 15
AAG26270
ID AAG26270 standard; protein; 342 AA.
XX
XX AAG26270;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 30665.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EF1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX
XX 05-MAR-1999; 99US-0123180P.
XX
XX 09-MAR-1999; 99US-0123548P.
XX
XX 23-MAR-1999; 99US-0125788P.

```

PR 25-MAR-1999; 99US-0136264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 07-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 19-MAY-1999; 99US-0134768P.  
PR 20-MAY-1999; 99US-0134941P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144844P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145132P.  
PR 22-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0147936P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.

Tue Apr 20 10:32:13 2004

```
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 15.0%; Score 476; DB 3; Length 342;
Best Local Similarity 36.6%; Pred. No. 3.3e-41;
Matches 113; Conservative 51; Mismatches 135; Indels 10; Gaps 5;

QY 32 VLDCVIPYIDDPKDRDAVSQVCRWYELDSLTKKHVTIALCYTTTPARLRRRPPHLESUK 91
Db 13 VLEHVFSFIQLDKRNSVSLVCKSWYERWCRCRVFIGNCYAVSPATVIRRPKVRVSE 72

QY 92 LKGKPRAMFNLIPEWGGHVTWPVXEISOYFCLXSLHFRMIVKXDSLQNLARDRHV 151
Db 73 LKGKPHFADFNLVPDGGWGYVPEIAMSSTYTWLEERILKRMVWTDCLLEIAKSFKN- 131

QY 152 LHALKDKCGFTTDLGHIGRCKSLRVLFLEESSILEKDGWHLALNNVTLETNF 211
Db 132 FKVLVLSCEGFTDGLAIAATCRNLKELDLRESDDVSGHWSHFPDPTYSLSVSLNI 191

QY 212 YLTDIAVVKIEDLELAKNCPNLVSKLTDCEILD-LVNPFFKHASALEPFCGGTYNEE-- 268
Db 192 SCL-ASEVSPSALERLVTRCPNLKSLNRAVPLEKLAITLQAPOLEELGTGGYTAERV 250

QY 269 PERYSALSLPAKLCR-----LGLTYICKNELPIVEMFAAVLKLDLLYAMLDTEHCLMI 323
Db 251 PDYISGLSVALSGCKELRCLSGFWDAVPAYLPVIVSVCSRLLTINLSYATVQSYDLVKLL 310

QY 324 QRCNLEVL 332
Db 311 CQCPKLQRL 319
```

Search completed: April 20, 2004, 01:36:33  
Job time : 80 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 01:34:16 ; Search time 25 seconds  
(without alignments)

1251.414 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180

Sequence: 1 TKTSAPFLFLSLRSNTTEE.....QRSDPDTVPVPLDTATCVDTI 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164.5	5.2	456	3	US-09-172-841-51
2	164.5	5.2	456	4	US-08-951-621-51
3	127.5	4.0	1151	4	US-09-177-165A-31
4	119	3.7	483	4	US-09-904-615-154
5	113.5	3.6	380	4	US-09-172-841-53
6	113.5	3.6	380	4	US-08-951-621-53
7	112.5	3.5	435	2	US-08-531-439B-4
8	111.5	3.5	1874	4	US-09-331-403-2
9	110.5	3.5	579	4	US-09-325-932A-185
10	109	3.4	1495	3	US-08-462-467B-12
11	108	3.4	870	4	US-09-252-991A-28407
12	104	3.3	1074	4	US-09-004-838-111
13	103.5	3.3	980	2	US-08-473-553A-6
14	103	3.2	985	2	US-08-473-553A-2
15	102	3.2	1257	3	US-09-081-823-5
16	102	3.2	582	4	US-09-081-149-7
17	102	3.2	777	1	US-08-526-964-3
18	102	3.2	777	2	US-08-946-617-3
19	102	3.2	777	3	US-09-031-897-3
20	101.5	3.2	456	2	US-08-910-731-2
21	101.5	3.2	456	2	US-08-795-395-2
22	101.5	3.2	1817	4	US-09-004-838-125
23	101	3.2	461	2	US-08-910-731-6
24	100.5	3.2	264	4	US-09-543-681A-8068
25	100.5	3.2	1323	4	US-09-004-838-90
26	100	3.1	1255	3	US-08-947-823-3
27	98.5	3.1	478	4	US-09-004-838-51

28	98.5	3.1	885	2	US-08-310-912A-2
29	98.5	3.1	885	3	US-08-841-089-2
30	98.5	3.1	885	3	US-09-301-085-2
31	98.5	3.1	885	5	PCT-US95-04570-2
32	98.5	3.1	885	5	PCT-US95-04589-2
33	98	3.1	409	4	US-09-252-991A-29674
34	98	3.1	922	4	US-09-328-352-5283
35	98	3.1	1206	4	US-09-245-928A-19
36	98	3.1	1243	4	US-09-245-928A-17
37	98	3.1	1247	4	US-09-803-286A-2
38	98	3.1	1257	4	US-09-245-928A-18
39	97.5	3.1	320	4	US-09-325-932A-190
40	97	3.1	456	2	US-08-910-731-8
41	97	3.1	777	4	US-09-257-770-7
42	97	3.1	1568	3	US-09-181-706-2
43	97	3.1	1568	3	US-09-458-791-2
44	97	3.1	1568	3	US-09-459-066-2
45	97	3.1	1568	4	US-09-459-065-2

#### ALIGNMENTS

#### RESULT 1

US-09-172-841-51  
; Sequence 51, Application US/09172841  
; Patent No. 6232081  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeffrey W.  
; APPLICANT: Elledge, Stephen J.  
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES  
; FILE REFERENCE: BCM-03510  
; CURRENT APPLICATION NUMBER: US/09/172.841  
; EARLIER FILING DATE: 1998-10-15  
; EARLIER APPLICATION NUMBER: 08/951,621  
; EARLIER FILING DATE: 1997-10-16  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 51  
; LENGTH: 456  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-172-841-51

Query Match 5.2%; Score 164.5; DB 3; Length 456;

Best Local Similarity 23.8%; Pred. No. 2e-09;

Matches 76; Conservative 51; Mismatches 124; Indels 69; Gaps 12;

QY	233	NLVSVKLTDCETLDLVNF-----FKHASALEEFCGGTYNEEPEP-----YSAISLPAKL	281
DB	45	NILALDGSNWQRIDLFNFQIDVEGRVVRNISKRCGGFLRKLSLRCIGVGDSSILKTFQON	104
QY	282	CB-----LGTITGKNELPVFMFAAVLKLDLYAMLDTEDHGM-LIQRCPNLEVL	333
DB	105	CNRIEHLNNGCKTIDTSTCVSLSPFCKLKHLDLTSCVINSILKISGECNLEYLN	164
QY	334	TR--NVIGDRGLEVGRCCRKRLRIERGDDOQMEDEEGTVSHRGLIALSQGSELEY	391
DB	165	LSWCQITIKGIEALVRGCGKAL-----LLRGCTQLED	199
QY	392	MAVYSDITNASLEHIGTHLKNLCDFRVLVLDHE--EKITDPLONGVPAALLRGCDKLAR	449
DB	200	EAL-----KHQNYC-HELVSINLQSCSRIT-----DEGVVQICRGCHRLQA	240
QY	450	FALYLRGGLTDVGLGYGYQSPNVWRMLLVVGE--SDAGLLEFAKGCPSLQKLENGCL	508
DB	241	LCL--SGCSNLTDSITAUGLNCPRIQLIARCSHLTDAGFTLLAENCHELEKMDLEECI	299
QY	509	FFSERALAAVAATQATSLRYL	528
DB	300	LITDSTLIQLSIHCPIKQAL	319

US-08-951-621-51  
US-08-951-621-51  
Sequence 51, Application US/08951621  
Patent No. 6573094  
GENERAL INFORMATION:  
APPLICANT: HARPER, JEFFREY W.  
APPLICANT: ELLEGE, STEPHEN J.  
TITLE OF INVENTION: F-BOX GENES AND PROTEINS  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Medlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/951,621  
FILING DATE: 16-OCT-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: MacKnight, Kamrin T.  
REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: BCM-02999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 456 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-951-621-51

Query Match 5.2%; Score 164.5; DB 4; Length 456;  
Best Local Similarity 23.8%; Pred. No. 2e-09;  
Matches 76; Conservative 51; Mismatches 124; Indels 69; Gaps 12;

QY 233 NLVSVKLTDCIEDLVNF-----FKHASALEEEFCGGTYNEEPER-----YSAISLPAKL 281  
DB 45 NILALDGSNWQRIDLFNFQIDVGRVVENISKRCGGFLRLSLRGCGIGVGDLSLKTFAQN 104  
QY 282 CR-----LGLTVIGKNELPIVFMFAAVLKLDLYAMLDTHCHM--LIQRCPNLEVL 333  
DB 105 CRNIEHLNNGCKTITDSTCYSLSRFCSKLKHLDLTSCVITNSSLKGISEGCRNLEYLN 164  
QY 334 TR--NVIGDRGLEVLGRCCXKRLKRLRIERGDDDDQGMDEEGTVSHRGLIALSQSCSELEY 391  
DB 165 LSWCDQITKDGIEALVRCGRGLKAL-----LIRGCTQLED 199  
QY 392 MAVTVSDITNASLEHIGTHLKNLDCDFRLVLLDHE--EKITDPLDNGVRALLGCDKLR 449  
DB 200 EAL-----KHQNYC--HELVSUNLQSCSKIT-----DEGVQICRCHRLQA 240  
QY 450 FALYLRRLGGTDLVGLVIGQYSPNVRWMLLYGYGE--SDAGLLEPAKGCPSLQKLEMRGCL 508  
DB 241 LCL--SGCSNLTASLTALGNLCPRLQILEARCSHLTDAGFTLLARNCHELEKMDLEECI 299  
QY 509 FFSERALAVALATQLSRYL 528  
DB 300 LITDSTLIQLSIHCPKLOAL 319

RESULT 3  
US-09-177-165A-31  
Sequence 31, Application US/09177165A

Patent No. 6426205  
GENERAL INFORMATION:  
APPLICANT: Tyers, Mike  
APPLICANT: Willem, Andrew  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN  
TITLE OF INVENTION: DEPENDENT PROTEOLYSIS  
FILE REFERENCE: 11757.10USUI  
CURRENT APPLICATION NUMBER: US/09/177,165A  
CURRENT FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: 60/092,443  
PRIOR FILING DATE: 1998-07-10  
PRIOR APPLICATION NUMBER: 60/063,254  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 31  
LENGTH: 1151  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-177-165A-31

Query Match 4.0%; Score 127.5; DB 4; Length 1151;  
Best Local Similarity 20.7%; Pred. No. 0.00018;  
Matches 119; Conservative 78; Mismatches 190; Indels 189; Gaps 29;

QY 11 LSLRSNMTEERNVRKTRVVVDVDCVIPY-----IDDPKDRDAV----- 49  
DB 212 MELRALETENMEIRNLRL--KILTIIEYKKSLEYACHSKLGRQQOVENFTDNFIWINSI 269  
QY 50 -----SQVCRWVELDSLTRKHVTIAL-----CYTTTPARLRRRFP----- 85  
DB 270 DTTESSDLKEGLQDSRYSRQFINNVLSNPSNQNTCTSVT-----RRSPFALNMLPSEI 324  
QY 86 -HLESILKLGKPRAMFNLPEDWGHVTPWVKEISQYEDCLKSLHFRMIKVSQDLQNL 144  
DB 325 LHLILDKLNQYDIYKFLTVSKLWA-----EI-----IVKILYRPHINKKSQDLDF 371  
QY 145 ARDRGHVHLAKLKDCKSGTTDGLPHIGRFGCKSLRVLFLEESSILEKGEWHELALNNT 204  
DB 372 -----LRTMKLT-----SEETVFNYRLMIRLNFSFV-----GDYMDHTELYNF 410  
QY 205 VLETLNFIYLTDAVVKIEDELLAKNCPNLVSKLTDCEILDVNFVKHASALEEFCGGT 264  
DB 411 V-----GCKNL-----ERLTIV-FCKHITSV----- 430  
QY 265 YNEEPERYSAISLPAKLRLGLTYIGKNELPIVFMFAAVLKLDLYAMLDTHCHM--L 322  
DB 431 -----PISAVLRGCKF-----LQSVDTIT-GIRDVSDVDVFTL 461  
QY 323 IQRCPNLE--VLETRNVIGD--RGLVLRGCCXKRLKRLRIERGDDDDQGMDEEGTVSHR 377  
DB 462 ATYCPRVQGYFVQARNVTDFSLRNFVHS---PMLKRIKITANNR---MNDE----- 508  
QY 378 GLI-ALSQSCSELEYNAVYVS-DITNASLEHIGTHLKNLDCDFRLVLLDHEEKITDLPDN 435  
DB 509 -LVELLANKCPLLVEVDITLSPNVTDSSLLKULTRVLQVLEFRIT---HMTNIT---DN 560  
QY 436 GVRALLRGCDKLRFPALYLRRG--GLTVDGLVIGQYSPNVRWMLLYGYGE--SDAGLLEF 492  
DB 561 LFQELSKVVDMPSLRLDLSGCENITDKTIESIVNLAFLKRLNVRFLGKCSRITDASLFOL 620  
QY 493 AKCPSLQKLEMRGCLFFSERALAVAATQLSRYL 528  
DB 621 SKLGNLQTVFHGHCNFTDNGVRALFHSCTRIOYV 656

RESULT 4  
US-09-904-615-154  
Sequence 154, Application US/09904615  
Patent No. 6566325  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 49 Human Secreted Proteins

```

; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; EARLIER FILING DATE: 2001-07-16
; SOFTWARE: Patent In Ver. 2.0
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 154
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (194)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-904-615-154

Query Match
Best Local Similarity 3.7%; Score 119; DB 4; Length 483;
Matches 96; Conservative 56; Mismatches 150; Indels 140; Gaps 19;

QY 90 LKLKGPRAAMNFI-IPEDW-----GGHVTWPVKELISQYFDCLSLHFR 132
Db 111 LSLVGRP-AGLSLRVQEPWADRARVLSLELVECAQASGVTEISITSETQQLCVQ-LEFP 168
QY 133 RMIVKDSLQNLARDRGHVLHAKLDKCSGTTDGLF--HIGRFCKSLRVLFLESSILE 190
Db 169 R-----QSENEPAVALRIAHCDLGAHSHLLXGQLMETCARLXQLSQVNICE 216
QY 191 KDGE---WLHELALNTVLETFNLYLTDIAVVKIEDLELLAK---NCPNLVSVKLTDCI 244
Db 217 DDDASSLLQSLLSLEKLT--FLTS-SCVSTEGLAHLASGLCHHL-----EE 265
QY 245 LDLVNFFKHASALEBFCGTYNEPERYSALSLPAKLCRLGLTYIGKNELPIVFMFAVL 304
Db 266 LDLSN-----NQDEEGTKALMRALEGKW-----ML 291
QY 305 KKLDLLYAMLTEDHCLMIOQPCNLEVLET-----RNVTGDRGLEVLGRCKELKELRIER 360
Db 292 KRLDLSHLLNSSTLALLTHRLSQMTCLQSLRLNRNSIGDVG-----CCHLSEALR--- 342
QY 361 GDDQGMEDDEGTSHRGLIALSQGCSELEYMAVYVSDITNASLEHIGTHLKNLCDFRLV 420
Db 343 -----RATSEELDLSHNOIGDAGVQHILATILPGLPELRKI 378
QY 421 LLDHE-----EKITDPLDNGVRALLRGCDKRLRFALYLRRLGLTDVGLGYIGQYSP 472
Db 379 DLSGNSISSAGGVQLAESLVLCRLLEMLGCNAL-----GDPTALGLAQ--ELFQ 427
QY 473 NVRWMLLVGVSADAGLLFAK 494
Db 428 HLRVHLHPFSLHGPFGALSLAR 449

RESULT 5
US-09-172-841-53
; Sequence 53, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; APPLICANT: Elledge, Stephen J.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT APPLICATION NUMBER: US/09/172,841
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621

```

```

; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 53
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-172-841-53

Query Match
Best Local Similarity 3.6%; Score 113.5; DB 3; Length 380;
Matches 97; Conservative 65; Mismatches 131; Indels 175; Gaps 24;

QY 57 YELDSLRKHVTIA-ICYTTPARLR-----RPHLESLLKLGKPRAMENLIPEDW 108
Db 48 YELIQLILNHLTPDLCLRLAQTCKLSHQHCDPLOYIHL-----NLQY-- 91
QY 109 GGHVTPWVKELISQYFDCLSLHFRFMIVKDSLQNLARDRGHVLHAKLDKCSGTTDGL 168
Db 92 -----WAK-----LDDTSLEFL-QSRCTLVQWNLN---SWTGNRGF 123
QY 169 FHIGRFCKSLRVLFLESSILEKDGWELHELALNTVLETFNLYLTDIAVVKIEDLELLA 228
Db 124 ISVAGFSRFLKVCSELVEL-----ELSCSHFLNETC-----LEVIS 160
QY 229 KNCNLVSVKLTDCIHLDLNFPFKHASALEBFCGTYNEPERYSALSLPAKLC---RLG 285
Db 161 EMCNQLQALNLSSCDKL-----PPQAFNHI---AKGCSLKKLV 195
QY 286 L--TYIGKNELPIVFMFAVLKLDLLYAMLTEDHCLMIOQPCNLEVLETRNVIQDRGL 343
Db 196 LYRTKVSQTALLSILNFCSELQHLSL-----GSCVMI-----EDYDVA--- 234
QY 344 EVLGRCKRLKRLRIERGGDDDDQGMEDDEGTSHRGLIALSQGCSELEYMAVYVSDITNAS 403
Db 235 SMIGAKCKKLTDLWRCKN-----ITENGIAELASGCPLEELDLGWCPTLOSS 284
QY 404 -----LEHIGTHLKNL-----CD-----FRVLDDHEEKITDPLDNGVRA 439
Db 285 TGCFTLHAQLPNLQKLFITANRSVCDTIDELACNCTRLQQLDLGLKVTI-----YKF 338
QY 440 LIRGC--DKLRRFALYLRRLGLTDVGLGYIGQYSPNVRWMLLVGVS 485
Db 339 VLNVCFDDRKANLRLVFRKKIFGYNKNFI-----LIRW--LGLIGNA 379

RESULT 6
US-08-951-621-53
; Sequence 53, Application US/08951621
; Patent No. 6573094
; GENERAL INFORMATION:
; APPLICANT: HARPER, JEFFREY W.
; APPLICANT: ELLEDGE, STEPHEN J.
; TITLE OF INVENTION: F-BOX GENES AND PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,621
; FILING DATE: 16-OCT-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.

```



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP97/07309  
FILING DATE: 29-DEC-1997  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1874 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-331-403-2

Query Match 3.5%; Score 111.5; DB 4; Length 1874;  
Best Local Similarity 22.7%; Pred. No. 0.029;  
Matches 107; Conservative 71; Mismatches 160; Indels 133; Gaps 27;

QY 79 RLRRFPHLSLKKGKPRAAFNLIPEDDGCHVTPWVKESQYFDCIKLSLHFRMIVKO 138  
DB 721 RLPRGFDKLSQLL-DISSKNFNIYEVINSCN--LLQLDLSYNKIRSL-----PD 770  
QY 139 S--DLQNLAR---DRGHVILHAKDKCSGFTTDLGLFHIGRF-----CK--SLRVLFLE 184  
DB 771 SMNQIQLAKINLSNNRITHVNDLSKMTSLRTDL---RYNRIESIKCRVFNQLNFLT 826  
QY 185 ES--SILEKGEWHLALANNVLETL---NFYLTDAIVVKIEDLEL-----LAKNCPN 233  
DB 827 ENRLTMFDDQLMLRTLELQRPNSILTKNDYLEHLTSLSKAKGAVLPESLRLRPR 886  
QY 234 LVSVKLTDCETILDVNFVKHASALEEFCGGTWNEE--PERSAI-----SL 277  
DB 887 LEKLELSNSITVLPDPIKHLKLVHLSVAKKLESDEIASLKNLAKMLDLHCNNMLT 946  
QY 278 PAKLCRLGLTYI-----GKNEL-----PIVFMFAA-----302  
DB 947 PAALSTLSITFNWISSNMLSGHHELYRTFQGTSTIAKSLMFLSAADNQMGKFWEIFNTE 1006  
QY 303 -VLKKDLIYAMLTDEHCLMIOQPNLEV--LETRNVIGRGLVLCRCCKRLRLRI- 358  
DB 1007 KTLKVLNLSY-----NNFMAL---PELEMENLTLYLSGNHLLTTSGEAFIKLSRLVL 1057  
QY 359 -ERGDGDDQMEDEGTVSHRGIALSQGCESELEY-MAVVVSDIT---NASLEHI-----407  
DB 1058 MLNANNLQSLPAEISQLSQ--LSVIDVGSNQLKYNLSNVHYDMNRQNTKYLNFSGNT 1115  
QY 408 -----GT--HLKNLCDFR-----LVLLDHEBKITDLPIDNGVRLLR 442  
DB 1116 RPEIKSAMIEYGTNNHLSDLTVLKQLRVGLGMDVTLNTRVP--DDGVNFRLR 1165

RESULT 9  
US-09-325-932A-185  
Sequence 185, Application US/09325932A  
Patent No. 6451604  
GENERAL INFORMATION:  
APPLICANT: Flimm, Barry  
TITLE OF INVENTION: Compositions affecting programmed cell  
TITLE OF INVENTION: death and their use in the modification of forestry plant develop  
FILE REFERENCE: 1022  
CURRENT APPLICATION NUMBER: US/09/325,932A  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 206  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 185  
LENGTH: 579  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-09-325-932A-185

Query Match 3.5%; Score 110.5; DB 4; Length 579;  
Best Local Similarity 22.0%; Pred. No. 0.0046;  
Matches 108; Conservative 70; Mismatches 158; Indels 155; Gaps 30;

QY 152 LHALKLDKCSGFTTDLGLFHIGRF-----CKSLRVLFLEBSILE 190  
DB 73 LSRLEILLRGESFVESIFSAGEIRQLVYLQWKECPISISFTIPTRNLVLYIQYALKT 132  
QY 191 KGEWHLALANNVLETLNFY-----LTDI--AVVKIEDLELA-----KNCN-----233  
DB 133 L---WQHE---SQAPLQTELYIDATLSRVPSQIGKLNQLERIVLKNKYFKTLFNEFYDM 186  
QY 234 --LVSVKLTDCETILDVNFVKHASALEEFCGGTWNEEPERYSALS-----LPKLCRL--GL 286  
DB 187 HSLKHITLQCEQOML-----LPDSVGLTGRQTHDFSGCSNQLQALPDSVGQLTGL 237  
QY 287 TYIKNE-----LPIVFMFAAVLKLDLYAMLTDEHCLMIOQPNLEVLETRNVIGD 340  
DB 238 KTLDELETSIQLPDSVGQLTGLQSLDL-----EHCTSLQLP-----DSVGQ 281  
QY 341 -RGLVIG-RCCKKLRLRIERGGDDQMEDEGTVSHRGIAL-SQGCSELEYMAVVVS 397  
DB 282 LTGLQTLDLRGCSL-----QGLPDSVGQLT--GLEGLYLSGCFSLQGLPDSVE 328  
QY 398 DITNA-----SLSHI-----GTHLKNL--CDFRLVLLDHEBKITDLPIDNGVR 438  
DB 329 QLTGLEGLYLSGCFSLQGLPDSVGQLTGLQSLINLEYCTSLGLEPDSVGQLTDLPLD---385  
QY 439 ALLRGCDKLRRFALYLRGGGLTDVGLGYI--GOYSPNVRMMLLYGVGSDAGLLEFAKGP 497  
DB 386 --INTCISLQ-----GLPD--SVGQLRGQLNLDLRWC-----DSLQGLPDSVGQLT 427  
QY 498 SLOKLEWRGCLFFGERALAVAAATQTSIRYLVWOG-----YGVSPSGRDLVWARPFM---550  
DB 428 GLQILDLSGCT--SLOGLPDSVGQLTGLRTLHLENCTSLQGLPDSVGNILTSK---WLNL 482  
QY 551 ----NIELIPS 557  
DB 483 SGCSNQLMPLN 493

RESULT 10  
US-08-462-467B-12  
Sequence 12, Application US/08462467B  
Patent No. 6210899  
GENERAL INFORMATION:  
APPLICANT: Rosenbaum, Jan S  
TITLE OF INVENTION: The Use of a BMP Protein Receptor  
TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells  
TITLE OF INVENTION: Co-Transfected with a Type II BMP Receptor and a Type I  
TITLE OF INVENTION: BMP Receptor  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Procter & Gamble Company  
STREET: 11810 East Miami River Road  
CITY: Ross  
STATE: OH  
COUNTRY: USA  
ZIP: 45061  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,467B  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Herbsko, Bart S.  
REGISTRATION NUMBER: 32,572  
REFERENCE/DOCKET NUMBER: 5474R  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (513) 627-0633  
TELEFAX: (513) 627-0260  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 1495 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-462-467B-12

```

Query Match 3.4%; Score 109; DB 3; Length 1495;  
Best Local Similarity 22.3%; Pred. No. 0.038;  
Matches 107; Conservative 44; Mismatches 164; Indels 16

167	QY	GLFHGRFOCKSRVLFLESSILEKDGCEWHELANNVTLETNFVLTDI	----	AVVKI	221
772	Db	GLTHRGLLILEYRGLNTHRVALLEMETARGHISGLASNILEBGLPHEILELALAAAS	831		
222	QY	EDELELA	-----KNCP-----NLVSVKLTDCETDLNVFPKHSALAEFCGGTY	265	
832	Db	PILELYSGCIYTHRGLYSERTPRTHRGINLETYREILETHRASPTYRYSGLAS	-----NGLY	889	
266	QY	NBEPERYAISLPAKLCRIGLTVIGKNELPVEMFAVLKLDLLVAMLDTEHCMLIQR	325		
890	Db	SERLETYRA-SPPEHL	-----ELYSCYSALATH	916	
326	QY	CPNILEVLEPNVIGDRGLEVLGRCKKELKELRIERGDDQDEEGTVSHRGLIALSQG	385		
917	Db	--PLEASPTHRARGALAELELYSLEALATYR	-----SERALALAA	955	
386	QY	CSELEYMAVTVSDITWASLEHIGTHKNLCDFRLVLLDHEKITDPLDNGVRALLRGCD	445		
956	Db	C---YSGLYLECYSHISLEHISTH	-----	976	
446	QY	KLRRFALYLRGGGLTVDGLGYIGQYSNVRWML	-----LGYCESDAGLLEF	495	
977	Db	--RGLILETYRGLYTHRGMLGVLVYSPRALAILEALAHISARGASPLELYSERLYSASN	1034		
496	QY	CPSLQKLEMRGCL	-----FFSERAL-----AVAATOLTSRLRYLWQGVGV	535	
1035	Db	ILELEILELYSLYSASNGLYSERCYSCVSIIEALAAASPLEGL	-----YLEALAVALLYSPEHAS	1093	
536	QY	-----SPSGR	-----DLLVMARPFWMNIELIPSKVATNTNPDETIVVVEHPAHILAYYSIAG	586	
1094	Db	NSBRASPTHRASNGLVALASP	-----ILEPRLASNTHRARGVALGYTHRL	1145	

RESULT 11

RESULT 11  
US-09-252-991A-28407  
Sequence 28407, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 28407  
LENGTH: 870  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28407

Query Match	3.4%	Score 108;	DB 4;	Length 870;
Best Local Similarity	21.6%	Pred. No. 0.019;		
Matches 140; Conservative	79;	Mismatches 240;	Indels 190;	Gaps 35;
QY	77	PARLRRRPHL-----ESLKUKGKPRAAFMFLIPEDMGWHTPMWKEISQVFDCLKS	128	
	88	PATFRRRLAGALGIPESDEACLCRLRNASALWLLDDYPRHPD---AFDACLCCLLS	144	

```

129 QY -----LHFRMIVKSDLIQ-----NLARDGRHVHLAKDKCQSGFTTGD 167
145 Db AASPRVGVWMIASRRRPPCNLRLLLEGELEVDDGGLLAFADEVAELLRL---HGSRADA 201
168 QY -----LFHTGRFCSLRVFLF--EESSIIKEKGEWLHELAINNTVLTETLNFYLDIAV 219
202 Db AATGALLERSGGWCALRLRLVGEEA----GPLLQEY-LQHEDDELPPPLADAARA 255
220 QY -----KIBDL-----ELAKCPNLSVK----- 238
256 Db LAWLPOVGPELFRLEPHELGHDLLARGPPLADANQRVSLPPAIRREAWREPQRFOR 315
239 QY ---LTDCEILDVNFFKHSALEBFCCGTYNBEPERYSAISLPKALC-----RLGIT 287
316 Db DFHREACEW----FAHGETREAVDQALAADPEF--TAAGLLOKLTEEQLLHGHNGMV 368
288 QY YIGKNELPIVMFEA----AVLKLDLYA--MLDTEDHCM--LIQPCNVLEIETRNVIG 339
369 Db LAURDELPAALLASTPRLVILNAWTLLYAGRLAEAE-D-CIGLARFLFPSPSARQVRILLA 427
340 QY D-RGL-EVLGRCCXKKRLRIERGDDD-----QGME-D-BEGTVSHRGILIALSQGCSE 388
428 Db QWOGLFGILLHC-----RGERGAADYIRLEALBOLPEDAWSOGICRSALMQLALIEGR 480
389 QY LEYMAVVYSBITNASLHGHGTHKNLCDF-RLVILDHEEKI--TDLPIDNGVYRAL----- 440
481 Db MDQARLIGRDALRLAREHDSLIFALIELEAQMLEQRBELLRAESVLDRAQRYLEDLQG 540
441 QY ----LRGCDKLLRRFALYLRGGLTDVGLGY---IQOYSPNV-RMWLLGVHGSDDAGLLEF 492
541 Db QGSPMLGRIALRRARLCIQQGREVEAGHWYRLGLEQARENLDPMALYGYL---LALLEA 597
493 QY AKG-----CPSLOKLE-----MRGCLPFSEERAJAAVAACLTLSRYLVWQYGV 535
598 Db QGGDLDAAFNRILLEVERLMQQRHVPDPYLRGALLIVSALTIO-----QSR 643
536 QY SPSGEDLLVMARPFWNIELISR-KVAINTNPDETWWVEHPAHILAYYS 583
644 Db PAQAREILLRWAYFQ----PGRARLSPPEPELEARVHQALAELYS 688

RESULT 12
US-09-004-838-111
; Sequence 111 Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
; APPLICANT: Michelmore, Richard W.
; APPLICANT: Shen, Kathy
; APPLICANT: Meyers, Blake
; TITLE OF INVENTION: Procedures and Materials for
; TITLE OF INVENTION: Conferring Pest Resistance in Plants
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,838
; FILING DATE: 09-JAN-1998
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/781,734
; FILING DATE: 10-JAN-1997
; ATTORNEY/AGENT INFORMATION:

```



```

NAME: Eihorn, Gregory P.
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 023070-078810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 1074 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: -
LOCATION: 1..1074
OTHER INFORMATION: /note= "RG2K deduced sequence"
US-09-004-838-111

```

Query Match	3.3%;	Score 104;	DB 4;	Length 1074;
Best Local Similarity	20.6%;	Pred. No. 0.078;		
Matches	91;	Conservative	124;	Indels 186; Gaps 21;
QY	136	VKSDLONLARDGRGHVLHAKLKDKC-----SGFTTDGLF-----HIGRECKSLRV	180	
DB	374	IEHHDIETIA---HVVFMQSYDNLQNEEAQSIFLLCGLPEDFDIPTTELVRVYGMGLRV	429	
QY	181	LFLEESSILKDGWMLHELAUNTVLETFNYL--TDIAVVKIBEDLLAKNCENLVSVK	238	
DB	430	F-----NGVVTIGEARHLNAYIELLKDSNLLIESDVHCKMHD-----	469	
QY	239	LTDCEILDVLNFFKHASALBEEFCGTYNEEPEERSAISLPACLK--RLGLTYIGKNELPIV	297	
DB	470	LVRAFVLDTNRFKHSLLVNHGNGMLG-WPEN---DMSASSCKRISLICKGMSDFPRD	524	
QY	298	FMFAAVL-----KKLDLL-----YAMLDTEHDC-----MLIQ	324	
DB	525	VKFPNLLILKLMHADKSLKFPQDFYGMKKLQVTSYDHMKYPLLPFTSPQCSTNLRVLHLH	584	
QY	325	RCP-----NLEVLETRNVIGDRGLVEVIGRCCKLRKLRIERGDDQGMEDBEG	372	
DB	585	QCSLMPDCSSIGNLLNLEVISFAN-----SGIEWLPSTIGNLKELR-----	625	
QY	373	TVSHRGUIALSQGCGSELEYMAVYVSDTDNASLEHIGTHLKNLCPDFRLVLDHBEKIYDLP	432	
DB	626	-----VLDLTN-----CD-----GLR	636	
QY	433	LDNGVRALLRGCDKLRRFALYLRGGTLTDVGLGVIGQYSPNVRWMLLGYVGESDAGLLEF	492	
DB	637	IDNGV---LKKLVKLEE--LYMRVGRYQKAI SF-----TDENCNEM	673	
QY	493	AKGCPISLOKLEMRGCLFFSEEA	514	
DB	674	AERSKNLSALEFE---FFKNA	692	

RESULT 13  
US-08-473-553A-6  
; Sequence 6, Application US/08473553A  
; Patent No. 5859338  
; GENERAL INFORMATION:  
; APPLICANT: Meyerowitz, Elliot M.  
; APPLICANT: Clark, Steven E.  
; APPLICANT: Williams, Robert W.  
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,  
; TITLE OF INVENTION: Transformed Plants, and Proteins  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States

```

; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 980 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-473-553A-6

Query Match          3.3%; Score 103.5; DB 2; Length 980;
Best Local Similarity 19.6%; Pred. No. 0.075;
Matches 117; Conservative 77; Mismatches 215; Indels 187; Gaps 28;

QY      9 FTLSLNSMTEERNVNRTRVDVLDCVIPIYIDDPKDRDAVSQCRRWYELDSLTKRH-- 66
Db      ||| : | : | : | : | : | : | : | : | : | : | : | : | : | :
349 FTQLPANLRGNGLIKLDYSNHLTGLI-----PKD-----LC-RGEKLEMLILSNFF 396
QY      67 -----VTIALCVTTTPARLRPPHLESCLKGKPRAAMFNL----- 103
Db      : | : | : | : | : | : | : | : | : | : | : | : | : | :
397 FFGIPPELGKKSLTKIRIVKNL-----LNGTVPAGLENPLVITIELTDNFSGEL 449
QY      104 -----IPEDW-GGHYTPWKKEISVFDCIKSLHFRRMIVKDSLOMLARDRG 149
Db      - - - - - IPEDW-GGHYTPWKKEISVFDCIKSLHFRRMIVKDSLOMLARDRG 149
450 PVTMSGDVDQLYLNNWFESGEIPPAIGN---FPNLQTFLDNRRFRG----NIPREIF 501
QY      150 HVHALKLKCSGETTGDLFHGRFCKSRLVLFEESSILEKGEWLHELANNVTLETLL 209
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
502 ELKHLSRINTSANNITGGIPDSISRCSTLISVDLSRNRI-----NGEIPKGI 548
QY      210 NFVLTDIADVKIIEDEL--LAKNCPNLVSKLFDCEILDVN-----FFKHASAL 257
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
549 N-NVKNLGTINISGNQUTGSIGPTGIGNNTSLTLDFENDLSGRVPLGGQFLVNETS-- 605
QY      258 EEFCGGYNEEPERYSAISLPACL-----CRGLGTYIG----- 290
Db      606 --FAGNYLCLPHRVSCPTRPGQTSDEHNHTALFPSRSIVITVAIAITGLILISVAIROMN 663
QY      291 --KNELPIVMFAAVLKDLLLYAMLDTEDHCMLIQPCNPLEVLETNRNVIGORGLEVLGR 348
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
664 KKKNQKSLAWKLTA-FOKLD-----FKSED-----VLECKEENIIKGGSGGVYR 708
QY      349 CCK-----RKRLRIERGD-DDQGMEDE-----EGTVSHRGLIALSQSCSELEYMAVVVS 397
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
709 GSMENNVDVAIKRL-VQRGTGRSDHGFTAETQTLGRIFHRHI VRLLVGVANKDTNLIYE 767
QY      398 DITNASLEHI-----GTHL-----KNLC-----DFRLVLDHEKITDLPDIN 435
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
768 YMPNGSIGELLHGSKGCHLQWETRHRVAVEAAKGLCYLHHDCSPLLIHRDVKSNILLDS 827
QY      436 GVRAALLRGCDKLRRFAFYLRGGTIDVGLGYGOYSPNVRMWLLGYGESDAGLLE 491
Db      : : : : : : : : : : : : : : : : : : : : : : : : : :
828 DFEAHVADFG---LAKFLVDGAASECMSSISTAGSY-----GYAPEAYATUK 870

```

RESULT 14

us-10-009-791-22.ra1

Tue Apr 20 10:32:14 2004

```

US-08-473-553A-2
; Sequence 2, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-473-553A-2

Query Match 3.2%; Score 103; DB 2; Length 985;
Best Local Similarity 19.9%; Pred. No. 0.086;
Matches 108; Conservative 71; Mismatches 191; Indels 174; Gaps 26;

QY 9 FTLSRNMTEERNVVRKTRVVDVLDVPIYDDPKDRDAVSQVCRWVLDLTKRH-- 66
Db 354 FTLQPLPAMLRGNLKLVDSDNHLTGLI-----PKD-----LC-RGEKLEMLLSNPF 401
QY 67 -----VTIALCYTTTPARLRPPHLESKLKGPRAAMFNL----- 103
Db 402 FPGPIPEELGCKSKTKIRIVKNL-----LNGTVPAGLFNLPLVTLIELTDNFFSGEL 454
QY 104 -----IPEDW-GGHVTPWVKESIQYFCLKSLHPRRMIVKSDIQLNLRDGR 149
Db 455 PVTMSGDVLDOIYLSNNWFSGEIPPAIGN-----FENLOTFLDRNRFRG-----NIPREIF 506
QY 150 HVLHAKLDCSGFTTDLGLHGRCKSLRVLFLEBSILEKDGWELHETLNTVLETL 209
Db 507 ELKHLSRINTSANNITGGIPDSISRCSTLISVDLSNRRI-----NGEIPKGI 553
QY 210 NFVLTDAIVVKEDLEL-----LAKNCNPLNSVKLTDCIEILDVN-----FFKHASAL 257
Db 554 N-NVKNLGTLNLSGNLTGSIPTGIGNMTSLTDLSDNLSGRVPLGGQFLVFNETS-- 610
QY 258 EEFCCGTYNEEPERYSAISLPAKL-----CRGLGYIG----- 290
Db 611 --FAGNTYLCPLHRVSCPTPGQTSDBHNTALFSPSRIVITVIAATGLILISVATROWN 668
QY 291 --KNEPLVFMFAVLKLLDLLVAMLDTEDHCLMIQRCNPLEVETRNVIQDGLGVLR 348
Db 669 KKNQKSLAWKLA-FQKLD-----FKSED-----VLECLKEENIIKGGSGIVYR 713

```

```

QY 349 CCK-----RLKRLRIERGD--DDQMEDE---EGTVSHRGLIALSQSCSELEYMAVYVS 397
Db 714 GSPFNVDVAIKRL-VORGTGRSDHGTAEIQTGLRPHRHIVRLGVVANKDTNLLYE 772
QY 398 DITNASLEHI-----GTHL-----KNLC-----DFRLVLDHEKITDPLDN 435
Db 773 YMPNGSLGELLHGSGKGLQWETRHRVAVEAAKGLCYLHHDSPILHRDVKSNLLDS 832
QY 436 GVRA 439
Db 833 DFEA 836

RESULT 15
US-08-947-823-5
; Sequence 5, Application US/08947823
; Patent No. 6114605
; GENERAL INFORMATION:
; APPLICANT: Williamson, Valerie M.
; APPLICANT: Kaloshian, Isagouhi
; APPLICANT: Yaghoobi, Jafar
; APPLICANT: Bodeau, John
; APPLICANT: Milligan, Stephen
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Pest Resistance in Plants
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/947,823
; FILING DATE: 09-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/18802
; FILING DATE: 09-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,191
; FILING DATE: 10-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-070210US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-947-823-5

Query Match 3.2%; Score 103; DB 3; Length 1257;
Best Local Similarity 21.0%; Pred. No. 0.13;
Matches 107; Conservative 67; Mismatches 176; Indels 160; Gaps 25;

QY 7 FLFTLSLRNMTEERNVVRKTRVVDVLDVPIYDDPKDRDAVSQVCRWVLEL-DSLTFK 65
Db 788 YLFTVYLG-----EGFVETKEMKGE-EVVKIYMDLISLSIV--IC--FNEIGDILNFQ 838
QY 66 HVTIALCYTTTPARLRPPHLESKLKGPRAAMFNLIPEDWGGHVTWVKESIQYFDC 125

```

Db	839	IHDVHDFCLIKAKENLFDRIK	-----SAPDLLPR-----QITIDYDEEEHF--	884
Qy	126	LKSLHFRMIVKDSDLQNLARDRGHVLHAKL	--DKCSGFTTDGLPHIGRCKSLRVLFL	183
Db	885	--GLNF--VMFDS--NKRHSGKHLYSLRINGDQLDDSVSDA	-FHL-RHLRLIRVLDL	934
Qy	184	ESSILEKDGEMWHELALNN	-----	203
Db	935	EPSLIMVND-SLLNEICMLNHLRYLIRITQVKYLPFPFSNLWNDESLFVSNKGSILVLLP		993
Qy	204	TVLETNLFYLDIAVVKIEDLE	-----LLAKNCPNLVSKLTDCEIL-----DLVN	249
Db	994	RILDVVKRLVLSVGACSFDDADESILIAKD	-----TKLENRILGELLISYSKDTMN	1047
Qy	250	FFKHASALEEFC	-----GGTYNEBPERYSALPAKLCRLGLTYIGN-----EL	294
Db	1048	IFKRPNLQVLQFELKESWDYSTQHWPFPKLDCLTELETCVGFKNNTNHCSSVVTNR		1107
Qy	295	PVFMFAAVLKLDLKYAMLTEDHCMLIQPCNLEVLETRNVIGDRGLEVLGRCKKELX		354
Db	1108	PWDFHPPSNLKL-LLYDFPLTSDSLSTIARLPNLENLSLYDTI	-----	1150
Qy	355	RLRIERGDDQGMED	-----BEGTVSHRGLIALS-QGCSELE-----	390
Db	1151	---IQGEWNNGEEDTFENLXFLNRLTLTKWVGEESFPNLEKLOECGKLEIIPPS		1207
Qy	391	---YMAVYVSDITNASLEHIGTHLKNLCD		416
Db	1208	FGDIYSLKFIKIVKSPQLEDSALKIKKYAE		1237

Search completed: April 20, 2004, 01:39:25  
Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

COM protein - protein search, using sw model

Run on: April 20, 2004, 01:38:26 ; Search time 362 Seconds  
(without alignments)

(without alignments)  
461.486 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180  
Sequence: 1 TKTSAFLETLRLRNMTTEE.....QRSDFPDVVPLDTATCVDT 606

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched. 1124875 seqs, 275673149 residues

total number of hits satisfying chosen parameters: 1124875

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Published Applications\_AA:

- 1: /cgm2\_6/ptodata/2/pubpaa/PCT\_NEW PUB. pep.\*
- 2: /cgm2\_6/ptodata/2/pubpaa/US06\_NEW PUB. pep.\*
- 3: /cgm2\_6/ptodata/2/pubpaa/US06\_PUBCOMB. pep.\*
- 4: /cgm2\_6/ptodata/2/pubpaa/US07\_NEW PUB. pep.\*
- 5: /cgm2\_6/ptodata/2/pubpaa/US07\_PUBCOMB. pep.\*
- 6: /cgm2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB. pep.\*
- 7: /cgm2\_6/ptodata/2/pubpaa/US08\_NEW PUB. pep.\*
- 8: /cgm2\_6/ptodata/2/pubpaa/US08\_PUBCOMB. pep.\*
- 9: /cgm2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB. pep.\*
- 10: /cgm2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB. pep.\*
- 11: /cgm2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB. pep.\*
- 12: /cgm2\_6/ptodata/2/pubpaa/US09\_NEW PUB. pep.\*
- 13: /cgm2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB. pep.\*
- 14: /cgm2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB. pep.\*
- 15: /cgm2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB. pep.\*
- 16: /cgm2\_6/ptodata/2/pubpaa/US10\_NEW PUB. pep.\*
- 17: /cgm2\_6/ptodata/2/pubpaa/US60\_NEW PUB. pep.\*
- 18: /cgm2\_6/ptodata/2/pubpaa/US60\_PUBCOMB. pep.\*
- 19: /cgm2\_6/ptodata/2/pubpaa/US60\_PUBCOMB. pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3112	97.9	592	12	US-10-425-114-45837	Sequence 45837, A
2	3095	97.3	590	12	US-10-424-599-284666	Sequence 284666, A
3	1786.5	56.2	598	15	US-10-310-154-520	Sequence 520, App
4	1701.5	53.5	591	15	US-10-310-154-519	Sequence 519, App
5	1584	49.8	313	12	US-10-425-114-49508	Sequence 49508, A
6	1305.5	41.1	425	12	US-10-425-114-58666	Sequence 58666, A
7	1275.5	40.1	410	12	US-10-425-114-69709	Sequence 69709, A
8	1193	37.5	394	12	US-10-425-114-46788	Sequence 46788, A
9	759.5	23.9	572	12	US-10-424-599-245726	Sequence 245726, A
10	737	23.2	587	12	US-10-424-599-234595	Sequence 234595, A
11	735	23.1	578	12	US-10-424-599-213007	Sequence 213007, A
12	693	21.8	218	12	US-10-424-599-143159	Sequence 143159, A
13	690.5	21.7	597	12	US-10-425-114-68189	Sequence 68189, A
14	676.5	21.3	598	15	US-10-310-154-655	Sequence 655, App
15	555.5	20.6	550	12	US-10-424-599-268553	Sequence 268553, A

16	646	20.3	152	12	US-10-424-599-268552
17	590.5	18.6	142	12	US-10-424-599-187454
18	485.5	14.0	462	12	US-10-424-599-231199
19	408	12.8	130	12	US-10-424-599-164173
20	369	11.6	321	12	US-10-424-599-240587
21	332	10.6	320	12	US-10-424-599-147777
22	332	10.2	162	12	US-10-424-599-177505
23	323.5	10.2	162	12	US-10-424-599-235954
24	323	10.2	162	12	US-10-424-599-222826
25	303.5	9.5	142	12	US-10-424-599-148467
26	303	9.5	204	12	US-10-424-599-182348
27	280	8.8	96	12	US-10-424-599-206714
28	277.5	8.7	596	12	US-10-424-599-147774
29	268.5	8.4	433	12	US-10-424-599-147774
30	242.5	7.6	660	15	US-10-310-154-521
31	234	7.4	432	12	US-10-425-114-59363
32	218	6.9	168	12	US-10-425-114-44523
33	202	6.4	315	12	US-10-425-114-55836
34	192	6.0	721	9	US-09-727-801-14
35	189.5	6.0	679	12	US-10-424-599-174551
36	188	5.9	179	12	US-10-424-599-361317
37	187.5	5.9	628	9	US-10-424-599-147775
38	184	5.8	105	12	US-09-727-801-21
39	184	5.8	639	12	US-10-424-599-177499
40	183	5.8	563	15	US-10-424-599-228553
41	179	5.6	422	15	US-10-310-154-657
42	179	5.6	422	15	US-10-028-248A-28
43	179	5.6	422	15	US-10-028-248A-86
44	179	5.6	422	15	US-10-107-782-28
45	179	5.6	422	15	US-10-107-782-86
46	179	5.6	436	15	US-10-435-696-91

## ALIGNMENTS

## RESULT 1

```

US-10-425-114-45837
; Sequence 45837, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45837
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700963303_FLI.pep
US-10-425-114-45837

```

	Query Match	97.9%	Score 3112;	DB 12;	Length 592;
	Best Local Similarity	100.0%;	Pred. No. 1.5e-307;		
	Matches 592;	Conservative	0;	Mismatches	0;
				Indels	Gaps
QY	15	SNMTEERNVNRKTRVVDVLDVPIIDIDPKORDAVSQVCRRWYELDSLTRKHVTI	ALCYT	74	
DB	1	SNMTEERNVNRKTRVVDVLDVPIIDIDPKORDAVSQVCRRWYELDSLTRKHVTI	ALCYT	60	
QY	75	TTPARLRRRFPHLESILKLGKPRAMFNLI	PEDWGGHVT	PWVKEI	SQYEDCLKSLHFRM
DB	61	TTPARLRRRFPHLESILKLGKPRAMFNLI	PEDWGGHVT	PWVKEI	SQYEDCLKSLHFRM
					120
					194

Db	121	IVKSDQLQNRDARGHVLHAKLKDCKSGFTTDLGPHIGRFGKSRVLPLESSILEKDCG	180
Qy	195	WLHELALANTVLETINFLYLTDAVVKIEDLELAKNCNLSVVKLTDCCEILDVNFPPKHA	254
Db	181	WLHELALANTVLETINFLYLTDAVVKIEDLELAKNCNLSVVKLTDCCEILDVNFPPKHA	240
Qy	255	SALEBFCGGTYNEEPERYSAISLPKLCRLGLTYIGKNELPIVFMFAAVLKKLDLLYAML	314
Db	241	SALEBFCGGTYNEEPERYSAISLPKLCRLGLTYIGKNELPIVFMFAAVLKKLDLLYAML	300
Qy	315	DTEDHMLIQRCPNLEVLTRNVIQDGLVLRGCKRLKRLRIERGDDQDQMEDEGTV	374
Db	301	DTEDHMLIQRCPNLEVLTRNVIQDGLVLRGCKRLKRLRIERGDDQDQMEDEGTV	360
Qy	375	SHRGLIALSQGCSELEYMAVYVSDITNASLEHIGTHLKNLCPRLVLLDHEEKITDLPD	434
Db	361	SHRGLIALSQGCSELEYMAVYVSDITNASLEHIGTHLKNLCPRLVLLDHEEKITDLPD	420
Qy	435	NGVRALLRGCDKLRFPALYLRGGTLDVGLGYIGQVSPNVRMMLGYGVGSDAGLLEFAK	494
Db	421	NGVRALLRGCDKLRFPALYLRGGTLDVGLGYIGQVSPNVRMMLGYGVGSDAGLLEFAK	480
Qy	495	GCPSLQKLEMRGCLFFSERALAAVAATQLTSLRYLWVQGYGVSPSGRDLVMAFPWNIEL	554
Db	481	GCPSLQKLEMRGCLFFSERALAAVAATQLTSLRYLWVQGYGVSPSGRDLVMAFPWNIEL	540
Qy	555	IPSRKVATNTNPDETIVVVEHPAHILAYYSLAGORSDFPDVVPDTCVDT	606
Db	541	IPSRKVATNTNPDETIVVVEHPAHILAYYSLAGORSDFPDVVPDTCVDT	592
RESULT 2			
US-10-424-599-284666			
; Sequence 284666, Application US/10424599			
; Publication No. US20040031072A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa Thomas J			
; APPLICANT: Kovalic David K			
; APPLICANT: Zhou Yihua			
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53223)B			
; CURRENT APPLICATION NUMBER: US/10/424,599			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 285684			
; SEQ ID NO 284666			
; LENGTH: 590			
; TYPE: PRT			
; ORGANISM: Glycine max			
; FEATURE:			
; OTHER INFORMATION: Clone ID: PAT_MRT3847_99079C.1.pep			
US-10-424-599-284666			
Query Match 97.3%; Score 3095; DB 12; Length 590;			
Best Local Similarity 99.5%; Pred. No. 8.3e-306;			
Matches 587; Conservative 3; Mismatches 0; Indels 0; Gaps 0;			
Qy	17	MTEERNVRKTRVVDVLDVCIPIYIDDPKDRDAVSQVCRWYELDSLTKKHVTIACVTTT	76
Db	1	MTEERNVRKTRVVDVLDVCIPIYIDDPKDRDAVSQVCRWYELDSLTKKHVTIACVTTT	60
Qy	77	PARLRPPHLESKLKGPAAAFNLIPEDWGGHVTWPWKEISQYDCLKSLHFRMIV	136
Db	61	PARLRPPHLESKLKGPAAAFNLIPEDWGGHVTWPWKEISQYDCLKSLHFRMIV	120
Qy	137	KDSDLQNLARDRGHVLHAKLKDCKSGFTTDLGPHIGRFGKSRVLPLESSILEKDCGWL	196
Db	121	KDSDLQNLARDRGHVLHAKLKDCKSGFTTDLGPHIGRFGKSRVLPLESSILEKDCGWL	180
Qy	197	HELALNTVLETINFLYLTDAVVKIEDLELAKNCNLSVVKLTDCCEILDVNFPPKHA	256

APPLICANT: Zeng, Xiaoping  
APPLICANT: Zhang, Qiang  
APPLICANT: Zhao, Yajuan  
APPLICANT: Zhou, Li  
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants  
FILE REFERENCE: 38-15(52796)B  
CURRENT APPLICATION NUMBER: US/10/310,154  
CURRENT FILING DATE: 2002-12-04  
PRIOR APPLICATION NUMBER: 60/337,358  
PRIOR FILING DATE: 2001-12-04  
NUMBER OF SEQ ID NOS: 736  
SEQ ID NO 520  
LENGTH: 598  
TYPE: PRT  
ORGANISM: Zea mays  
US-10-310-154-520

Query Match 56.2%; Score 1786.5; DB 15; Length 598;  
Best Local Similarity 60.1%; Pred. No. 1.7e-172;  
Matches 349; Conservative 85; Mismatches 136; Indels 11; Gaps 7;

QY 28 VDVVLDVIVYIDDPKDRDAVSQVCRWYELDSLTKRHVTIALCYTTTPARLRFPPL 87  
DB 22 VPEALHLVFGYMDPDRRAASVLCRUHRIKALSRKHVTGVCYAVEPARLLARFPL 81  
QY 88 ESKLKGKPRAMFNLIPEWGGHVTWPWKESIQVDFCLSKLHFRMIVKDSLDQNLARD 147  
DB 82 ESKLKGKPRAMFNLIPEWGGHVTWPWKESIQVDFCLSKLHFRMIVKDSLDQNLARD 141  
QY 148 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLPFLESSILEKDG-EWLHELALNTVL 207  
DB 142 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLPFLESSILEKDG-EWLHELALNTVL 201  
QY 208 TLFVLTIDIAVKIE--DLELLAKNCPNLVSKLTDCEILDVNFPHKASALEFCGGTY 265  
DB 202 TLFVLTIDIAVKIE--DLELLAKNCPNLVSKLTDCEILDVNFPHKASALEFCGGTY 258  
QY 266 NE--EPERYSAISLPAKLCRL-GLTYIGKNELPIVFMFAAVLKKLDLLAYMLDTEHCL 322  
DB 259 FEVGEYTKYKVIFFPRCLFGLTFMGKNEMPIFFYSTMLKKLDLQFTLTTEDHCL 318  
QY 323 IQRCNLEVTNRVIGRGLVGRVLCRCRKLRLRIEGRDDOGDEGTVSHRGLAL 382  
DB 319 IAKCNLSLVNRVIGRGLVGRVLCRCRKLRLRIEGRDDOGDEGTVSHRGLAL 377  
QY 383 SQGCELEYMAVYSDITNASLEHIGTGLKNCDFRLVLLDHEEKITDPLDNGVRALLR 442  
DB 378 AVGCELEYMAVYSDITNASLEHIGTGLKNCDFRLVLLDHEEKITDPLDNGVRALLR 437  
QY 443 GCDKLRRFALYLRGLTVDGLGYTQYSPNVRWMLLVGVGESDAGLLEFAKGCPSLKL 502  
DB 438 NCVKLRRAFAVLRPGGLSDVGLGYTQYSPNVRWMLLVGVGESDAGLLEFAKGCPSLKL 497  
QY 503 EMRGCLFESERALAAVQTLSTRLVWQGVSPSGRDLVWAPFNWELIPS-RKVA 561  
DB 498 ELRGCC-FESERALAAVQTLSTRLVWQGVSPSGRDLVWAPFNWELIPS-RKVA 556  
QY 562 TINTNPDETIVVHRPAHILAYYSLAGORSDFPVTVPVPLDTAT 602  
DB 557 YRVMDGQPCVTHAQVLAAYSLAGRRPDCQWLVTLHPAS 597

RESULT 4  
US-10-310-154-519  
Sequence 519, Application US/10310154  
Publication No. US2003023670A1  
GENERAL INFORMATION:  
APPLICANT: Edgerton, Michael D  
APPLICANT: Chomet, Paul S.  
APPLICANT: Adams, Thomas H  
APPLICANT: Ruff, Thomas G.  
APPLICANT: Agarwal, Aneta K.  
APPLICANT: Ahrens, Jeffrey E.

APPLICANT: Ball, James A.  
APPLICANT: Banu, G.  
APPLICANT: Bell, Erin  
APPLICANT: Boddupalli, Raghava  
APPLICANT: Deikman, Jill  
APPLICANT: Deng, Molian  
APPLICANT: Dong, Jinzhao  
APPLICANT: Duff, Stephen M.  
APPLICANT: Galligan, Meghan M.  
APPLICANT: Hinchey, Brenda S.  
APPLICANT: Huang, Shihshieh  
APPLICANT: Johnson, G. Richard  
APPLICANT: Jung, Vincent  
APPLICANT: Kretzmer, Keith A.  
APPLICANT: Laccetti, Lucille B.  
APPLICANT: Lai, Chao-Qiang  
APPLICANT: Lee, Gary  
APPLICANT: Lin, Jie-Yi  
APPLICANT: Liu, Jingdong  
APPLICANT: Lu, Bin  
APPLICANT: Luethy, Michael M.  
APPLICANT: Lund, Adrian  
APPLICANT: Madson, Linda L.  
APPLICANT: Malloy, Kathleen A.  
APPLICANT: McKiel, Christine L.  
APPLICANT: Miller, Philip W.  
APPLICANT: Padmavathi, Manchikanti  
APPLICANT: Parnell, Laurence D.  
APPLICANT: Start, William G.  
APPLICANT: Tennessee, Dan  
APPLICANT: Vidya, K.R.  
APPLICANT: Wang, Haiyuan  
APPLICANT: Xin, Zhaoguo  
APPLICANT: Xu, Nanfei  
APPLICANT: Yang, Chunzhi  
APPLICANT: Zeng, Xiaoping  
APPLICANT: Zhang, Qiang  
APPLICANT: Zhao, Yajuan  
APPLICANT: Zhou, Li  
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants  
FILE REFERENCE: 38-15(52796)B  
CURRENT APPLICATION NUMBER: US/10/310,154  
CURRENT FILING DATE: 2002-12-04  
PRIOR APPLICATION NUMBER: 60/337,358  
PRIOR FILING DATE: 2001-12-04  
NUMBER OF SEQ ID NOS: 736  
SEQ ID NO 519  
LENGTH: 591  
TYPE: PRT  
ORGANISM: Zea mays  
US-10-310-154-519

Query Match 53.5%; Score 1701.5; DB 15; Length 591;  
Best Local Similarity 57.0%; Pred. No. 7.7e-164;  
Matches 330; Conservative 95; Mismatches 135; Indels 19; Gaps 7;

QY 28 VDVVLDVIVYIDDPKDRDAVSQVCRWYELDSLTKRHVTIALCYTTTPARLRFPPL 87  
DB 20 IPDTALGLVGVYEDPDRDAISLVCRHWCVRDALSRKHVTVMAYSTTPELFRFPCL 79  
QY 88 ESKLKGKPRAMFNLIPEWGGHVTWPWKESIQVDFCLSKLHFRMIVKDSLDQNLARD 147  
DB 80 ESKLKGKPRAMFNLIPEWGGHVTWPWKESIQVDFCLSKLHFRMIVKDSLDQNLARD 139  
QY 148 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLPFLESSILEKDG-EWLHELALNTVL 206  
DB 140 RGHVHLAKDKCSGFTTDLGPHGRFCKSRVLPFLESSILEKDG-EWLHELALNTVL 199  
QY 207 ETLNPFYLDIAVKIELELLAKNCPNLVSKLTDCEILDVNFPHKASALEFCGGTY 266  
DB 200 ETLNPFYLDIAVKIELELLAKNCPNLVSKLTDCEILDVNFPHKASALEFCGGTY 258  
QY 267 BEPE-----RYSAISLPAKLCRLGLTYIGKNELPIVFMFAAVLKKLDLLAYMLDTEHCL 321

```

Db      259 EQGQPVASRNENYFFPSLHLSILYMGTDMLQILFFYATALKKLDLQFTFLSTEDHCQ 318
Qy      322 LIQCPNLVLETRNVIGDRGLVGRCKLRIRIERGDDQ--GMEDEEGTVSHRGLI 380
Db      319 IVQCSNLETLEVRDVIQDRGLVQVVAQCKLHRURVERGDDQOGLDEQGRISQVGLM 378
Qy      381 ALSQGCSELEYMAVVSDDITNASLEHIGTHLKNLCDFRLVLLDHEEKITDPLDNGVRAL 440
Db      379 AIAQCPELTWAIHVSDDITNALEAVGTCSKNLNDLFLVLLDREAHITELPLDNGVRAL 438
Qy      441 LRGCCKLRARFALYLRGGTVDGLGYIGQYSPNVRWMLLGVYGESDAGLLFAKGCPSLQ 500
Db      439 LRGCCKLRARFALYLRGGTVDGLGYIGQYSPNVRWMLLGVYGESDAGLLFAKGCPSLQ 498
Qy      501 KLEMPGCLFFSERALAVATOLTSRYLWVQGYGSPSGRDLVWARPFWNIELPSRKV 560
Db      499 KLEVRGCL-FSEHALALAAQLKLSRYLWVQGYGSPSGRDLVWARPFWNIELPSRKV 553
Qy      561 ATNTNPD-ETVVVEHPAHILAYISLAGQSDPDPDTVWPL 598
Db      554 -----PDQDEPCPEKHKQLAYISLAGRTDCPPSWTL 587

```

```

RESULT 5
US-10-425-114-49508
; Sequence 49508, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49508
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701002452_FLI.pcp
US-10-425-114-49508

```

```

Query Match      49.8%; Score 1584; DB 12; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.7e-152;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      305 KKLDLLYAMLDTHDCHMLIORCPNLEVTRNVIGDRGLVGRCKLRKLRIRIERGDD 364
Db      12 KKLDLLYAMLDTHDCHMLIORCPNLEVTRNVIGDRGLVGRCKLRKLRIRIERGDD 71
Qy      365 QGMEDEEGTVSHRGLIALSQGCSELEYMAVVSDDITNASLEHIGTHLKNLCDFRLVLLDH 424
Db      72 QGMEDEEGTVSHRGLIALSQGCSELEYMAVVSDDITNASLEHIGTHLKNLCDFRLVLLDH 131
Qy      425 EEKTTDPLDNGVRALLRGCDKLRFPALYLRGGTVDGLGYIGQYSPNVRWMLLGVYGE 484
Db      132 EEKTTDPLDNGVRALLRGCDKLRFPALYLRGGTVDGLGYIGQYSPNVRWMLLGVYGE 191
Qy      485 SDAGLLFAKGCPSIQKLEMRGCLFFSERALAVATOLTSRYLWVQGYGSPSGRDLV 544
Db      192 SDAGLLFAKGCPSIQKLEMRGCLFFSERALAVATOLTSRYLWVQGYGSPSGRDLV 251
Qy      545 MARFPFWNIELPSRKVATNTNPDVVEHPAHILAYISLAGQSDPDPDTVWPLDTATCV 604
Db      252 MARFPFWNIELPSRKVATNTNPDVVEHPAHILAYISLAGQSDPDPDTVWPLDTATCV 311

```

```

Qy      505 DT 606
Db      312 DT 313

```

```

RESULT 6
US-10-425-114-56666
; Sequence 56666, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56666
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73276H05_FLI.pcp
US-10-425-114-56666

```

```

Query Match      41.1%; Score 1305.5; DB 12; Length 425;
Best Local Similarity 61.0%; Pred. No. 1e-123;
Matches 261; Conservative 64; Mismatches 92; Indels 11; Gaps 7;

Qy      181 LFELESSILEKDGEMWHELAANTVLEINFLYTDIAVVKIE--DLELLAKNCPNLVSVK 238
Db      2 LFELECIIEDEGEWHELAANNVLTINFTYMTL---KVEPADLELLAKNCKSLISLK 58
Qy      239 LTDCEILDVNFVKHASALEEFCGGTYNE--SPERYSALSIPAKLCRL-GLTYIGKNELP 295
Db      59 MGDCLSDSLIGFPQTSKALQEFAGGAFVEGTYTKEKVIFFPRCLFCGLTFMGKNEMP 118
Qy      296 IVFMFAAALKLDLLAYMLDTHDCHMLIORCPNLEVTRNVIGDRGLVGRCKLRKLR 355
Db      119 VIFPSTMLKLDLQFTLTTHDCHQIAKCPNLVLEVRNVIGDRGLVAVATCKLR 178
Qy      356 LRIRGDDDDQGMDEBEGTVSHRGLIALSQGCSELEYMAVVSDDITNASLEHIGTHLKNLC 415
Db      179 LRIRGDDDDPGQE-EQGVSGIQLTAVAVGCRELEYIAAYVSDITNGALESIGTFCKNLY 237
Qy      416 DFLVLLDHEEKITDPLDNGVRALLRGCDKLRFPALYLRGGTVDGLGYIGQYSPNVR 475
Db      238 DFLVLLDQKKTIDPLDNGVRALLRNCVKLRFAFYLRPGSLSDVGLGYIGLYSGNIQ 297
Qy      476 WMLLVGVGESDAGLLFAKGCPSLQKLEMRGCLFFSERALAVATOLTSRYLWVQGYG 535
Db      298 YMLLVGVGESDAGLLFAKGCPSLQKLEMRGCLFFSERALAVAVLQVPSRYTVWQGYRA 356
Qy      536 SPSEGRDLVWARPFWNIELIPS-RKVATNTNPDVVEHPAHILAYISLAGQSDPDPDT 594
Db      357 SRTGQDMLMARPYWNIETFAPIPEISAYRVMDGQPCVDTHAQVLAAYISLAGRPPCPQW 416
Qy      595 VVPLDSTAT 602
Db      417 LVTLHPAS 424

```

```

RESULT 7
US-10-425-114-69709
; Sequence 69709, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

```



```

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69709
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17190F07_FLI.pep
; US-10-425-114-69709

Query Match 40.1%; Score 1275.5; DB 12; Length 410;
Best Local Similarity 61.4%; Pred. No. 1.1e-120;
Matches 254; Conservative 62; Mismatches 87; Indels 11; Gaps 7;

QY 195 WLHMLANNVTLEINLYDIADVVKIE--DELLAKNCNLYSVKLTDCIILLDNFFK 252
DB 1 WLHMLAVNNSVLKINFYMTL---KVEPADLELLAKNCKSLSLKMGDCDLSDLIGFFQ 57
QY 253 HASALEFECGTYNE--EPERYSAISLPKLCRL-GLTYIGKNELPVFMFAVLKKLDL 309
DB 58 TSXALQBFAGAGFAFVEGYEKYKVIIPPRLCFCFGLGUTFMGKNEMPVIFPYSTMLKKLDL 117
QY 310 LYAMLDTEHCHMLTQRCNLEVLERNVIGDRGLEVLRGRCCKRLRLRIERGDDDDQKMD 369
DB 118 QFTLTTHEDHQLAKCPNLSVLEVRNVIGDRGLEVAATCKLRLRIERGDDDDQGE- 176
QY 370 BEGTVSHRGLIALSQGSELEYMAVYSDITNASLEHIGTHLKNLCDFRLVLDHEKIT 429
DB 177 EOGVSGITGLTAVAVGCRELEYIAAYVSDITNGALESIGTFCKNLYDFRLVLDKOKKIT 236
QY 430 DPLDNGVRALLRGCDKLRRLVLRGGLTDLVGLYIGYSPNVNRMILGYGVEDAGL 489
DB 237 DPLDNGVRALLRNCVKLRFAFLRPGGLSDVGLYIGYSGNIQYMLLGNVGEDNGL 296
QY 490 LEFAGKCPESLOKLEMRGCLFFSERALAVAAVQTLTSLRYLWQYGVSPSGRDLVMPARPF 549
DB 297 IQFATGCTNLKLELRGCC-FSERALAVAVLQMPESLRYINWQYGRASRTGQDLMMPARPY 355
QY 550 WNIELI-PSRKVANTNPDETVVVEHPAHILAYYSLAGORSDFPDVTVPLDTAT 602
DB 356 WNIEFAPPSPESAYRVMAADQPCVDTHAQVLAAYYSLAGRRPDCQWLVTLHPAS 409

RESULT 8
US-10-425-114-46788
; Sequence 46788, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46788
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: 700800510_FLI.pep
; US-10-425-114-46788

Query Match 37.5%; Score 1193; DB 12; Length 394;
Best Local Similarity 61.7%; Pred. No. 2.6e-112;
Matches 234; Conservative 57; Mismatches 84; Indels 4; Gaps 3;

QY 223 DLELLAKNCNLYSVKLTDCIILLDNVNFPHASALEFECGTYNEPE--RYSAISLPK 280
DB 14 DLELLAKSCKSLTSLKIGDCDLSDLIGFFQSATSLEEFAGGTNGOGELTKYGDVFPSPR 73
QY 281 LCELGLTYIGKNELPVFMFAVLKKLDLYAMLDTEHCHMLTQRCNLEVLERNVIGD 340
DB 74 ICSLGLTFMGANEMPIIFFFSALKKLDLQYTFETTEDHCOLIAKCPNLLVAVRNVIGD 133
QY 341 RGLVLRGRCCKRLRLRIERGDDDDQKMDDEBEGTVSHRGLIALSQGSELEYMAVYSDIT 400
DB 134 RGLGVVADTCKLQRLRIERGDDDEGGVQEGVQVGLTAJAVGCRELEYIAAYVSDIT 193
QY 401 NASLEHIGTHLKNLCDFRLVLDHEKITDPLDNGVRALLRGCDKLRRLVLRGGLT 460
DB 194 NGALESTGTFCCKLYDFRLVLDREERITDPLDNGVRALLRGCTMLRRFALYLRPGGLS 253
QY 461 DVGLEYIGYSPNVNRMILGYGVEDAGLLEFAKGCPSLOKLEMRGCLFFSERALAVAA 520
DB 254 DAGLEYIGYQCSGNIQYMLLGNVGETDDGLISFALGCNLRKLELRSCC-FSERALALAIL 312
QY 521 QLTSRLYLWQYGVSPSGRDLVMPARPFWNIELI-PSRKVANTNPDETVVVEHPAHIL 579
DB 313 SMESLRYVWQYKASQTRDLMLMARPFWNIEFTPPSSQNAQRLIEDGFCVDSHAQIL 372
QY 580 AYSLAGORSDFPDVTVPL 598
DB 373 AYSLAGKRLDCQSVVTL 391

RESULT 9
US-10-424-599-245726
; Sequence 245726, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245726
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63922C.1.pep
; US-10-424-599-245726

Query Match 23.9%; Score 759.5; DB 12; Length 572;
Best Local Similarity 33.2%; Pred. No. 6.8e-68;
Matches 194; Conservative 95; Mismatches 256; Indels 39; Gaps 12;

QY 30 DWLDCVPIYDDPKRDVAVSOVCRRWYELDSLTRKHVTIALCYTTTPARLRPPHLES 89
DB 7 DEVIEHIFDVVVSHDRNALSLVCKSWYRIERCTQRVFIGNCYSITPERLIQRPGLKS 66
QY 90 LKLKGPAAFNILPEWDGWHVTPWKEISQYDCLKSLHFRMIVKDSLDONLARDG 149
DB 67 LTLKGPHFADFSLVYDNGGFWHPWIBALAKNVKGLEELRLKRMVVSDESLELSRSET 126
QY 150 HVLHAKLDKCSGGTTDGLFHIGRFCKSLRVLFLFESSILEKDGFWLHELANNVTLETL 209

```

```

Db      127 H-FKSLVLVCEGFSTDGLAALANCRFLRELDLOENEVEDEHKQWLSLSCFPDNCISLVS 185
QY      210 NFYLTDAVVKIEDLELLAKNCPNLVSVKLTDCETILD-LVNFFKHASALEEFCGGTYNEE 268
Db      186 NFACLK-GEVSLGALERFVARSPNLKSLKNSRSPVDALQRMRAPOQLSDLGIGSLVHD 244
QY      269 PERYSALSLPAKLCRL-----GLTYIGKNELPIVEMFAAVLKKLDLLYAM-LDTEHDC 320
Db      245 PESAYIKLNTILKCKSITSLSGFLVAPICLAAIYPICNLSLMSLYAAGIQGSDLV 304
QY      321 MLQRCNLEVLTRNVIGRGLVIGRCCKRLRLR-----IERGDDDDQGMEDREGTVSH 376
Db      305 KLIHCVKQLRWIMDCIGDKGLGVWATTCQDLQELAVFSPVPGD-----PAAVTE 356
QY      377 RGLTALSGCSELEYMAVYSDITNASLEHIGTHLKNLCDFRLVLLDHEKITDL--PLD 434
Db      357 KGLVAISMGCPCKLSLDFCHQMTNAALITVAKNCPNFIIRFLCILDATKPDPTMQPLD 416
QY      435 NGVRAILRGCDKLRFAIYLRGGITDVGIGYQYSPNVRWMLLGYVGBSDAGLLEFAK 494
Db      417 EGFCNIVQSCRLRLSL-----SGKLTQVFLYIGYAEKLEMLSTAFAGDGDKMGLYLN 473
QY      495 GCPSLQKLEMRGCLFFSERALAAVATQLTSLRYLVQGYGVSFSGRDLVYMARPFWMNIEL 554
Db      474 GCKLRLKLEIRDCP-FGNMALLTDVGKYETMRSLWSSCEVTGACKLLAKKHPRLNVEI 532
QY      555 IPSRKVATNTPDETIVVVEHPAHT---LAYYSLAGQSDPFDTV 595
Db      533 F-----NENEQBCDSLEGDQKVKMYLYRTLAGKDAPEYV 569

```

## RESULT 10

```

US-10-424-599-234595
; Sequence 234595, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234595
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(587)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53867C.1.pap
US-10-424-599-234595

```

```

Query Match      23.2%; Score 737; DB 12; Length 587;
Best Local Similarity 32.7%; Pred. No. 1.4e-65;
Matches 188; Conservative 108; Mismatches 253; Indels 26; Gaps 14;

QY      33 LDCVIPYIDDPKDRDAVSQVCRRWYELDSLTRKHVTIALCYTTTPARLRRRPFHLESKL 92
Db      15 LEVTFPMNERDRNAISLVCKSWYIERWCRSKVFGNCIYAVSPVLMVVRPEVTSIAL 74
QY      93 KGKPRAMFNLIPEDEGGHVTVPWKEISQYFDCLSLHFRFMRIVKDSLDQNLARDRGHVL 152
Db      75 KGKPFADFDLVPGGWCYICPWIDAVRSFPCFEQFLKTKWTWITDESLELAKSPFN-F 133
QY      153 HALKDKCSGFTTDGLFHIGRCFKSLRVLFLESSILEKD--GEMLHELANNVTLETIN 210

```

```

Db      134 KVLVLTSCGKXTPGLAAIAAANCRNLRELDLQRESEV-EEDLCGHWSLHSHFPDSTYLSVLN 192
QY      211 FYLTDAVVKIEDLELLAKNCPNLVSVKLTDCETILD-LVNFFKHASALEEFCGGTYNEE- 268
Db      193 ISCLN-NEVSLSALERLLGRCPNLRTLRANRAVPLDRPLNLLFAPQLVELGTGVYSTEM 251
QY      269 -PERYSALSLPAKLCR-----LGLTYIGKNELPIVEMFAAVLKKLDLLYAMLTDEHCL 322
Db      252 RPEVFNLEAAFAFGCKQLKSLSGFWDVLPSTLPAVVPICSRSLTSLNLSYAIIOSSDLIKL 311
QY      323 IQRCNLEVLTRNVIGRGLVIGRCCKRLRLRTERGGDDOGMEDEEGTVSHRGLIAL 382
Db      312 ISQCNLLRWLVLDYIEDAGLYAALASCKDLRELVR-FPSDDPFGLEPNV-ALTEQGGVSV 369
QY      383 SQGSELEYMAVYSDITNASLEHIGTHLKNLCDFRLVLLDHE--EKITDPLDNGVRL 440
Db      370 SEGCPKLSVLYFCRQMSNAALHTIARNRTNLTRFLCIIEPTPTDYLTHEPLDSGFGAI 429
QY      441 LRGCCKLRRLRYLRGGITDVGIGYQYSPNVRWMLLGYVGBSDAGLLEFAKGCPSLQ 500
Db      430 VEQCKLOQLSL---SGLLTDRVFEYIGTCGKKLEMLSVAFAGDSDLGLHHVLSGCDNLR 486
QY      501 KLEMRCGLFFSERALAAVATQLTSLRYLVQGYGVSFSGRDLVYMARPFWMNIELIPSKV 560
Db      487 KLEIRDCP-FGDKALLANAELKLETRSLWSSCEVSGACKLLGQKMPRLNVEVIDERG- 544
QY      561 ATNTPDETIVVVEHPAHTIAYYSLAGQSDPFDTV 595
Db      545 PDSRPDDCPVEK---LYIYRTIAGPRLDMPGFV 575

RESULT 11
US-10-424-599-212007
; Sequence 212007, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212007
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33469C.1.pap
US-10-424-599-212007

```

```

Query Match      23.1%; Score 735; DB 12; Length 578;
Best Local Similarity 32.0%; Pred. No. 2.2e-65;
Matches 187; Conservative 107; Mismatches 251; Indels 40; Gaps 13;

QY      30 DVVLDCVIPYIDDPKDRDAVSQVCRRWYELDSLTRKHVTIALCYTTTPARLRRRPFHLES 89
Db      18 DEVLERILGMLKSKDKSTVSLVCKEWFENAEWRSSRVFIGNCYSVSPILLTRFRFNTRS 77
QY      90 LKLCKPRAMFNLIPEDEGGHVTVPWKEISQYFDCLSLHFRFMRIVKDSLDQNLARDRG 149
Db      78 VTLGKPRFSDNLPANWAGADIHSMVAFKWPWLEELRKRMTVTDESLEFLAKFP 137
QY      150 HVLHALKDKCSGFTTDGLFHIGRCFKSLRVLFLESSILEKDGEMLHELANNVTLETI 209
Db      138 N-FKALLSLSCDGFSTDGLASATCKNLTELDIOENGIEDKSNLWSCFPDSFTSLEVL 196
QY      210 NF--YLTDAVVKIEDLELLAKNCPNLVSVKLTDCETILD-LVNFFKHASALEEFCGGTYN 266
Db      197 NFANLHND---VNFDALEKLVSRCKSLATLKNVKSVTLEQLQRLVHVHPQLGELGTGSFS 253

```

```

267 QY BE--PERYSATISLPAKLCR-----LGLTYIGKNELPIVFMEEAAVLKKLDLLLYAMLDTEDH 319
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
254 Db QELTSQCCSDLESALKCKNLHTLSGLWATAQIYLPVLSACTNLTFLNFVSAPLDSGDL 313
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

320 QY CMLTORCPNLEVLTRNVIGDRGLEVLGRCCKLXRLIERCDD--DOGMEDDEGVTVSHRG 378
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

314 Db AKLVHPCKQLORIWWVDVEDKGLEAVGSHCPLEEELRVFPADPFDEGIHV---GVTESG 370
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

379 QY LIALSQGSSELEYMAVVSDITNASLEHGTHLNKLCDFRLVL-----LDHEEKITDPL 433
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

371 Db FIAVSQGCPRLHYVLVFCROMTNAAVTVQCNPDFTHRCLIMHPGGLDY---LTQESM 427
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

434 QY DNGVRALLRGCDKLRFRFALYIRRGGLTVGLGYICQYSPNVKRWMLLYGVGESDAGILLEFA 493
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

428 Db DEAFGAVVKTCKLQRLAV---SGYLTDLTEYIGKYAKNLETLSVAFAGSSDWGRCVIL 484
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

494 QY KGCSLOKLWNRGCCLFFSERALAVAATOLTSURLYLWQGYGPSQRDLLVMARFPWNIE 553
    ||||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

485 Db DGCXPRLKLEVRDCP -FGNGALLSGLGKVESWRSLWMSDCNLTMTGVRLLAQEMPLNVE 543
    ||||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

554 QY LIPSRKVATNTNPDETUVVVEHPAHILAYYSLAGQRSDFPDVTVPVL 598
    ||||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

544 VI-----KEBSYETHOAKKVYVRSVAGPRRDAPPFFVILT 578
    ||||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

```

```

RESULT 12
US-10-424-599-143159
; Sequence 143159, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143159
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100286C.1.pep
US-10-424-599-143159

```

Query Match	21.8%	Score 693;	DB 12;	Length 218;
Best Local Similarity	77.6%;	Pred. No. 9e-62;		
Matches 132;	Conservative 14;	Mismatches 14;	Indels 10;	Gaps 2;
QY	4	SAPFLFTLSLSNNTEERNVKK--TRVVVDVLLDCVPIYIDDPKORDAVSQVCRRWYELD	60	
DB	49	NAPFLAVT-----EDRDARMTARRUSDVLLDCVMPYIHDSDKORDAVSQVCRRWYELD	101	
QY	61	SILTRKHVTIALCYTTTPARLRFRFPFLHSLSLKKGKPRAAFMNLIPEDWGGHVTTPVWKEIS	120	
DB	102	SILTRKHVTIALCYTTTPDLRFRFPFLHSLSLKKGKPRAAFMNLIPEDWGGFTTPWVRVIS	161	
QY	121	QYFDCCLKSLHFRFMIVKQSDLQNLARDRGHVLHALKDKCSGFTTDGLFPH	170	
DB	162	QYFDCCLKSLHFRFMIVRSDLQVLAARSRGHVLQALRLNCSGFTSDGLYY	211	

RESULT 13  
US-10-425-114-68189  
; Sequence 68189, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68189
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4570-007-All_F11.pep
US-10-425-114-68189

```

Query Match	21.7%	Score 690.5	DB 12	Length 597
Best Local Similarity	31.2%	Fred. No. 7.9e-61		
Matches	181	Conservative 114	Mismatches 251	Indels 35
Gaps	14			
Qy	32	VLDQVPIYDDKDRDAVSOVCRWYELDSLTKRHVTIALCYTTTPARLRFRFPHLSLK	91	
Db	30	VWEHFSFLSHSDRNTVSLVCKWYIEVERLSRAVFGVNCYAVRPERVVLRFENIKALT	89	
Qy	92	LKGKPRAAENLIPEDWGGCHVTWVKEISOYFCLKSLHFRMIKVDSDIQNLARD--RG	149	
Db	90	VKGKPHADFNLPVDPWGGYAGPWIDAARSCVGLEELRMKRMKVVFVDENLELLARSELRF	149	
Qy	150	HVLHALKDKSGFTTDLGPHIGFCKSLAVLEFESSILEKDGWEHELALNNVLETL	209	
Db	150	KV----LVLSICEGFTDLAAASHCKLLRELDQENDVDRGPRWL--SPFDSCTSLVSL	205	
Qy	210	NFYLTDTIAVKIE----DLELLAKNCNLVSVKLTDCIELD-LYNFFKHASALEEFCOGT	264	
Db	206	NF-----ACTIGEVNSGALERVARSFNLSRLNRSVSDTLLSKILLRAPNLEDLGTGN	260	
Qy	265	YNEE--PERYSAISLPKRLGLTYTGKE-----LPIVFMFAVLKLLDLLYA-MLDT	316	
Db	261	LTDEFOAESYRSLTSALECKKLSLSGFWDASPICVPIYPLCHQLTGLNLSYPTLUDY	320	
Qy	317	EDHMLQRCNLEVBLETRNVIGDRGLEVLRCGRCKRLKRLRIERGDDDDQGMDEBEGTVSH	376	
Db	321	SDLAKMVRCKLQRLWLVDCISDKGLQVASSCKDLQELRV---PPEFVVPVPGASAVTE	377	
Qy	377	RGLTALSQGCSELEYMAVYVSDITNASLEHGTGHLKMLCDPRVLLDHE--EKITDLDLD	434	
Db	378	EGLVAISGCPKLTSLLYFCHQMTNEALITVANNCPNFIREFLCILPEKPKEDAMTQGLD	437	
Qy	435	NGVTALLRGCDLRRFPALYLRGGTLTVGLGYGOYSPNVWMLLGYVGSDDAGLLEPAK	494	
Db	438	EGFGAIVRECKGLRRLSI---SGULTDKVFMVYIGKHAKYLEMLGISAFAGDSDKGMWVMN	494	
Qy	495	GCPSIQLEMRGCLFFSERALAVAATQUTSLRYLVWQYGYSPSGRDLILVMARPFWNL	554	
Db	495	GCKNLKLEIRDSP-FGDVALLGNVAKYETMRSLWMSSCNVTLKGCQVLASKMPMLNVEI	553	
Qy	555	IPSRKVATNTNPDETTVVEHPAHILAYYLAGQRSDPDDTV	595	
Db	554	M--NELDGGSEMEHNGDLSKVDKLYVYRTTAGARDADAFNV	592	

RESULT 14  
US-10-310-154-655  
; Sequence 655, Application US/10310154  
; Publication No. US2003023670A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; APPLICANT: Chomet, Paul S.  
; APPLICANT: Adams, Thomas H.  
; APPLICANT: Ruff, Thomas G.  
; APPLICANT: Agarwal, Ameeta K.

```
search completed: April 20, 2004, 01:49:29
Job time : 364 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 22:59:35 ; Search time 19 Seconds

(without alignments)  
1660.764 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180

Sequence: 1 TKTSAFFLTLSRSNWTEE.....QRSDFPDVVPLDTCVDT 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	183.5	5.8	461	1 YKK7 CAEEL	P34284 caenorhabdi
2	179	5.6	436	1 FL2L HUMAN	Q96192 homo sapien
3	175	5.5	423	1 FL2L MOUSE	Q8bh16 mus musculu
4	175	5.5	436	1 FL2L MOUSE	Q9czv8 mus musculu
5	172	5.4	491	1 FL17 HUMAN	Q9ujt9 homo sapien
6	170.5	5.4	423	1 FL2L HUMAN	Q9ukc9 homo sapien
7	154	4.8	463	1 POF2 SCHPO	O74783 schizosacch
8	147.5	4.6	276	1 FL2L RAT	Q9qzh7 rattus norv
9	141	4.4	539	1 FL16 HUMAN	Q8n531 homo sapien
10	139	4.4	630	1 ESAB TRYEQ	P26337 trypanosoma
11	133	4.2	630	1 ESAB TRYEB	P23799 trypanosoma
12	132	4.2	565	1 RAD7 YEAST	P06779 saccharomyc
13	131.5	4.1	1061	1 NA12 HUMAN	P39046 homo sapien
14	127.5	4.0	1151	1 GRR1 YEAST	P24814 saccharomyc
15	126.5	4.0	621	1 FL16 HUMAN	Q9uka2 homo sapien
16	120.5	3.8	535	1 FL16 MOUSE	Q9qxw0 mus musculu
17	119	3.7	1024	1 POPC BALSO	Q9rbs2 ralstonia s
18	118	3.7	699	1 YQOA CAEEL	Q09299 caenorhabdi
19	117.5	3.7	1093	1 NA14 HUMAN	Q86w24 homo sapien
20	115	3.6	434	1 FX3B HUMAN	Q9ukt6 homo sapien
21	114.5	3.6	434	1 FX3B MOUSE	Q8bfz4 mus musculu
22	110.5	3.5	326	1 FX1C HUMAN	Q9nxk8 homo sapien
23	110.5	3.5	424	1 SKP2 HUMAN	Q13309 homo sapien
24	109	3.4	424	1 FX1C MOUSE	Q9epx5 mus musculu
25	108	3.4	424	1 SKP2 MOUSE	Q92023 mus musculu
26	106.5	3.3	880	1 GYRA HAEIN	P43700 haemophilus
27	105.5	3.3	223	1 FX19 HUMAN	Q8wv35 homo sapien
28	105.5	3.3	1034	1 CLS1 HUMAN	Q96p20 homo sapien
29	104.5	3.3	980	1 CLV1 ARATH	O9syq8 arabidopsis
30	103	3.2	859	1 MUTS AQUAE	O66f52 aquifex aeo
31	103	3.2	868	1 MUTS XYLFT	Q87c18 xyellia fas
32	103	3.2	1039	1 YR71 CAEEL	Q09564 caenorhabdi
33	102.5	3.2	566	1 Y397 MYCGE	P47637 mycoplasma

## RESULT 1

ID	YKK7 CAEEL	STANDARD;	PRT;	461 AA.
AC	P34284;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Hypothetical F-box/LRR-repeat protein C02F5.7 in chromosome III.			
GN	C02F5.7.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Caenorhabditis.			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Bristol N2.			
RX	MEDLINE=94150718; PubMed=7906398;			
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton L., Dear S., Du Z., Durbin R., Favello A., Fraser A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laisater N., Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R., Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J., Wohldman P.;			
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans";			
RL	Nature 368:32-38(1994).			
RN	[2]			
RP	REVISIONS.			
RL	Waterston R.;			
CC	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: Contains 1 F-box domain.			
CC	-1- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; L14745; AAA27922.2; --			
DR	PIR; S44609; S44609			
DR	WormPep; C02F5.7; CE24780.			
DR	InterPro; IPR001810; F-box.			
DR	InterPro; IPR001611; LRR.			
DR	InterPro; IPR007089; LRR_CyB.			
DR	InterPro; IPR006553; LRR_CyB_sub.			
DR	Pfam; PF00646; F-box; 1.			
DR	Pfam; PF00560; LRR; 2.			
DR	SMART; SM00256; FBOX; 1.			

34	102.5	3.2	1024	1	CARC_HUMAN
35	102.5	3.2	1131	1	PHYA_SOYEN
36	102	3.2	582	1	SHO2_MOUSE
37	102	3.2	834	1	AIIM_YEAST
38	101.5	3.2	456	1	RINI_PIG
39	101	3.2	460	1	RINI_HUMAN
40	101	3.2	681	1	COG2_CAEEL
41	101	3.2	951	1	LGR4_HUMAN
42	100.5	3.2	2052	1	UBRB_SCHPO
43	100	3.1	446	1	AK_RICPR
44	100	3.1	2364	1	SPCO_HUMAN
45	99	3.1	399	1	R24L_ARATH

## ALIGNMENTS

Q9npp4 homo sapien  
P42500 glycine max  
O88520 mus musculu  
P03875 saccharomyc  
P10775 sus scrofa  
P13489 homo sapien  
Q21444 caenorhabdi  
Q9bxb1 homo sapien  
O13731 schizosacch  
Q9zc17 rickettsia  
Q01082 homo sapien  
Q9c646 arabidopsis

DR SMART; SM00367; LRR CC; 5.  
DR PROSITE; PS50181; FBOX; 1.  
KW Hypothetical protein; Ubl conjugation pathway; Repeat;  
KW Leucine-rich repeat.  
FT REPEAT 54 100 F-BOX.  
FT REPEAT 122 147 LRR 1.  
FT REPEAT 148 173 LRR 2.  
FT REPEAT 174 199 LRR 3.  
FT REPEAT 200 225 LRR 4.  
FT REPEAT 226 251 LRR 5.  
FT REPEAT 252 277 LRR 6.  
FT REPEAT 278 303 LRR 7.  
FT REPEAT 304 329 LRR 8.  
FT REPEAT 330 355 LRR 9.  
FT REPEAT 356 381 LRR 10.  
FT REPEAT 408 433 LRR 11.  
SQ SEQUENCE 461 AA; 51552 MW; BB28C98A5CA7B1D5 CRC64;  
  
Query Match 5.8%; Score 183.5; DB 1; Length 461;  
Best Local Similarity 22.0%; Pred. NO. 1.2e-06;  
Matches 98; Conservative 63; Mismatches 134; Indels 151; Gaps 17;  
  
QY 50 SQVCRVVEL-----DSLTR-----KHVTIALCVTTTPAR 79  
DB 81 AQVCKRSILADGSNQVRDLFTFQRDVKTAVVENLARRCGGLKSLKGCENVHDSA 140  
QY 80 LR---RRPHLESILKLGKPRAMPNLIPEDWGHVTFWVKISQYFDCLKSLHFRMIV 136  
DB 141 LRTFTSRCPNLEHLSL-----YRCKRV 162  
QY 137 KSDLQNLARDGHVHLKLDKSCGFTTDLGPHIGRCKSLRVLFLEESSILEKDGHWL 196  
DB 163 TDASCENLGR-YCHKLYNLENCSITDRAMKYIGDGCPLNSYLINISWCDAIQDGV-- 219  
QY 197 HELALNN-TVLETNFY---LTDIAVVKIDLELLAKNCPNLVSVKLTDCIEIDLNVFF 251  
DB 220 -QILSNCKSLDTLILRCCEGLTENFVGSVAHMAIKNLNLCQFLTITVQNIAN-- 276  
QY 252 KHASALEFCGGTYNEEPYSAISLPKLCRLGLTYIGKNELPIVFWFAVVKKLDLLY 311  
DB 277 -GATALEYLCMNCNQISDR-----SLVSLG----- 301  
QY 312 AMLDTEHCMLIQPCNLEVLTR--NVIGDGLVLCRCCKRLKRLRIERGDDDDQGMED 369  
DB 302 -----QSHNLIKVLSELSCTLIHGNGFTPLARGCRQLRLDME----- 339  
QY 370 BEGTVSHRGLIALSQCSSELEYMAV-YVSDITNASLEHIGT-HLKNLCDPRLVLLDHEEK 427  
DB 340 DCSLSIDHTINSANNCTALRELSLSHCELITDESIGNLASKHRETL--NVLELDNCPQ 396  
QY 428 ITDPLDNGVALLRGCDKLRFPALY 453  
DB 397 LTDSTLSH-----LRHCKALKRIDLY 417  
  
RESULT 2  
FL2L HUMAN STANDARD; PRT; 436 AA.  
AC Q96IG2;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DE 15-MAR-2004 (Rel. 43, Last annotation update)  
DE F-box/LRR-repeat protein 2-like.  
GN FBL2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Carninci P., Prange C.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Abramson R.D., Mullaly S.J.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Malek J.A., Gunaratne P.H.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Richards S., Worley K.C., Hale S., Sodergran E.J., Lu X., Gibbs R.A.,  
RA Villalon D.K., Muzny D.M., Sodergran E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalka U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -I- FUNCTION: Probably recognizes and binds to some phosphorylated  
CC proteins and promotes their ubiquitination and degradation.  
CC -I- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
CC -I- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -I- SIMILARITY: Contains 1 F-box domain.  
CC -I- SIMILARITY: Contains 12 leucine-rich (LRR) repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; BC007557; AA07557.2; -.  
CC InterPro; IPR001810; F-box.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR007089; LRR\_CYS.  
DR InterPro; IPR006553; LRR\_CYS\_sub.  
DR Pfam; PF00646; F-box; 1.  
DR Pfam; PF00560; LRR; 2.  
DR SMART; SM00256; FBOX; 1.  
DR SMART; SM00367; LRR CC; 12.  
DR PROSITE; PS50181; FBOX; 1.  
KW Ubl conjugation pathway; Leucine-rich repeat; Repeat.  
FT DOMAIN 22 68 F-BOX.  
FT REPEAT 90 115 LRR 1.  
FT REPEAT 116 141 LRR 2.  
FT REPEAT 142 167 LRR 3.  
FT REPEAT 168 193 LRR 4.  
FT REPEAT 194 219 LRR 5.  
FT REPEAT 220 245 LRR 6.  
FT REPEAT 246 271 LRR 7.  
FT REPEAT 272 297 LRR 8.  
FT REPEAT 298 323 LRR 9.  
FT REPEAT 324 349 LRR 10.  
FT REPEAT 353 377 LRR 11.  
FT REPEAT 378 403 LRR 12.  
SQ SEQUENCE 436 AA; 48423 MW; 39CD04A505C8CE3E CRC64;  
  
Query Match 5.6%; Score 179; DB 1; Length 436;  
Best Local Similarity 22.5%; Pred. NO. 2.5e-06;  
Matches 74; Conservative 47; Mismatches 114; Indels 94; Gaps 11;  
  
QY 151 VUHALDKSCGFTTDLGPHIGRCKSLRVLFLEESSILEKDGHWLHLMNNTVLETLN 210  
DB 170 LLEQLNISWCDDQVTKDGIQALVRGCGGLKALFLCKCTQLED----- 210  
QY 211 FYLTIDIAVVKIDLELLAKNCPNLVSVKLTDC-EILDVNFPMKASALEEFCGGTYNEEP 269  
DB 211 -----EALKYIGNHCPVLTNLNQLTQITD-----EGLTTCRGCHKLQ- 250





REPEAT	246	271	LRR 7.
REPEAT	272	297	LRR 8.
REPEAT	298	323	LRR 9.
REPEAT	324	349	LRR 10.
REPEAT	353	377	LRR 11.
REPEAT	378	403	LRR 12.
VARSPLIC	277	436	Missing (in isoform 2).
CONFLICT	40	40	/FTIQ-VSP 008968.
SEQUENCE	436 AA;	48390 MW;	L -> P (IN REF. 1; BAB28039).
			C80CE1861AF21BC3 CRC64;

Query Match  
Best Local Similarity 5.5%; Score 175; DB 1; Length 436;  
Matches 73; Conservative 47; Mismatches 115; Indels 94; Gaps 11;

QY	151	VLHALKDKCSGFTTDLGFHIGRFCKSLRVLFLEBSILEKGEWLHELALNNTVLETLN	210
DB	170	LLEQLNISCDQVTKDGIQALVRGCGLKALFLKGCTQLED-----	210
QY	211	FYLTDIAVKIEDLELLAKNCNPISVKLTDC-EILDLVNFFKHASALEEFCGGTYNEEP	269
DB	211	-----EALKYGAHCPELVTLNIQTCHITD-----EGLTIICRGCHKLQ-	250
QY	270	EYSATISLPALKRLCLGTIVYGKNELPIVFMAFVAALKLDLYLAMLPDEHMLIQRCPNL	329
DB	251	-----SLCASCSNITD-----AINALG-----QNCPRL	275
QY	330	EVLETR-NVIGDRGLEVLGRCCRKLRRIERGDDDDQGMEDEGTVSHRGLIALSOGCS	387
DB	276	RILEVARCSQLTDVGFTTLARNCHEKMDLE-----ECVQITDSTLIQLSHTCP	325
QY	388	ELEYMAV-YVSDDTNASLEHGTHLXNLCDFRLVLLDHBEKIITDLPIDNGVRALLRGCDK	446
DB	326	RQVLSSLSCHELTDDGIRHLGNACAHDLQLEVIELDNCPLIITDASLEH-----LKSCPS	380
QY	447	LRREFALYLRGGGLTDVGLGVGYQSYPNVR	475
DB	381	FERIELYDCQ-QITRAGIKELRTHLPNIK	408

### RESULT 5

FXL7 HUMAN	STANDARD;	PRT;	491 AA.
ID	-FXL7 HUMAN		
AC	Q9UT9;	O94926;	
DT	10-OCT-2003	(Rel. 42, Created)	
DT	10-OCT-2003	(Rel. 42, Last sequence update)	
DT	15-MAR-2004	(Rel. 43, Last annotation update)	
DE	F-box/LRR-repeat protein 7 (-Box and leucine-rich repeat protein 7).		
GN	FBXL7 OR FBXL7 OR FBXL7 OR KIA0840.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20399565; PubMed=10945468;		
RA	Ilyin G.P., Riialand M., Pigeon C., Guguen-Guillouzo C.;		
RA	"cDNA cloning and expression analysis of new members of the mammalian		
RT	F-box protein family";		
RL	Genomics 67:40-47(2000).		
RM	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=99156230; PubMed=10048485;		
RA	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,		
RA	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;		
RT	"Prediction of the coding sequences of unidentified human genes. XII.		
RT	The complete sequences of 100 new cDNA clones from brain which code		
RL	for large proteins in vitro.";		
RL	DNA Res. 5:355-364(1998).		
RN	[3]		
RP	SEQUENCE OF 41-493 FROM N.A.		
RX	MEDLINE=20003060; PubMed=10531035;		

	QY	402	ASLEHGTHGNKVLCDPRLVLLDHEEKITDPLDNGVRALLRGCDKLRRFALVLR-RG--G	450	-----RYLSIAHCGRVTDV-----GIRVVAKYCSKLR-----YINARGCEG	389
	Db	343	REIAKLESRL-----RYSIAHCGRVTDV-----GIRVVAKYCSKLR-----YINARGCEG	450	-----RYLSIAHCGRVTDV-----GIRVVAKYCSKLR-----YINARGCEG	389
	QY	459	LTPDVGLGYIGOYSPNVWMLLGVYG-ESDAGILBFAGKCPSLQKLEMRCGLFFSERALAV	517	-----RYLSIAHCGRVTDV-----GIRVVAKYCSKLR-----YINARGCEG	448
	Db	389	ITDHGVEYLAKNCTKLKSLOLGKPLVSDTGLECLALNCFNILKRSLSKSCESIITGOGLQI	448	-----RYLSIAHCGRVTDV-----GIRVVAKYCSKLR-----YINARGCEG	448
	QY	518	AATQLTSLRYLVWGYYGVSPSGRDLLVNARFWNIELIPSRKVATNTNP	566	-----RYLSIAHCGRVTDV-----GIRVVAKYCSKLR-----YINARGCEG	566
	Db	449	VAAACFDLQTLNVQDCVS-----YEALRF--VGKCKRCVIEHTNP	488	-----RYLSIAHCGRVTDV-----GIRVVAKYCSKLR-----YINARGCEG	488

RESULT 6

FL2_HUMAN	STANDARD;	PRT;	423 AA.
ID	Q9UKC9; Q9UK27; Q9UKA5; Q9Y3Y9;		
AD	Q9UKC9; Q9UK27; Q9UKA5; Q9Y3Y9;		
DT	15-MAR-2004 (Rel. 43, Created)		
DT	15-MAR-2004 (Rel. 43, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, last annotation update)		
DE	F-box/LRR-repeat protein 2 (F-box and leucine-rich repeat protein 2)		
DE	(F-box protein FBL2) (F-box protein FBL3).		
OS	FBL2 OR FBL2 OR FBL3.		
GN	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_Taxid=9606;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RX	MEDLINE=2003060; PubMed=10531035;		
RA	Cenciarelli C., Chiaur D.S., Guardavaccaro D., Parks W., Vidal M.,		
RA	Pagano M.;		
RT	"Identification of a family of human F-box proteins.";		
RL	Curr. Biol. 9:1177-1179(1999).		
RN	[2]		
RN	SEQUENCE FROM N.A.; INTERACTION WITH SKPL, AND TISSUE SPECIFICITY.		
RX	MEDLINE=20003061; PubMed=10531037;		
RA	Winston J.T., Kepp D.M., Zhu C., Elledge S.J., Harper J.W.;		
RA	"A family of mammalian F-box proteins.";		
RL	Curr. Biol. 9:1180-1182(1999).		
RN	[3]		
RN	SEQUENCE FROM N.A.		
RX	MEDLINE=20399565; PubMed=10945468;		
RA	Ilyin G.P., Riialand M., Pigeon C., Guguen-Guillouzo C.;		
RA	"cDNA cloning and expression analysis of new members of the mammalian		
RT	F-box protein family.";		
RL	Genomics 67:40-47(2000).		
RN	[4]		
RN	SEQUENCE FROM N.A.		
RA	Isoagi T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,		
RA	Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,		
RA	Tanase T., Nomura Y., Toshiya S., Komai F., Hara R., Takeuchi K.,		
RA	Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,		
RA	Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;		
RT	"NEDO human cDNA sequencing project.";		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.		
RN	[5]		
RN	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Honig L.,		
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Prange C.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Vallalon D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettenman M., Madan A., Rodrigues S., Sanchez A.,		

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RL human and mouse cDNA sequences.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
 [6]  
 RC SEQUENCE OF 174-423 FROM N.A.  
 RC TISSUE=Brain;  
 RA Wambutt R., Heubner D., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -!- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
 CC protein ligase complex (By similarity).  
 CC -!- TISSUE SPECIFICITY: Expressed in brain, heart, kidney, liver,  
 CC lung, pancreas and placenta.  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC  
 CC -----  
 DR EMBL; AF174589; AAF04510.1; -;  
 DR EMBL; AF176518; AAF03128.1; ALT\_INIT.  
 DR EMBL; AF186273; AAD56248.1; -;  
 DR EMBL; AK001438; BAA91691.1; -;  
 DR EMBL; BC031556; AAH31556.1; -;  
 DR EMBL; AL049953; CAB43222.1; -;  
 DR EMBL; HGNC:13598; FBX12.  
 DR MIM; 605652; -;  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR007089; LRR\_cys.  
 DR InterPro; IPR006553; LRR\_cys\_sub.  
 DR Pfam; PF00646; F-box; 1.  
 DR Pfam; PF00560; LRR; 3.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00367; LRR\_CC; 2.  
 DR PROSITE; PS0181; FBOX; 1.  
 KW Ub1 conjugation pathway; Leucine-rich repeat; Repeat.  
 FT DOMAIN 9 55  
 FT REPEAT 77 102  
 FT REPEAT 103 128  
 FT REPEAT 129 154  
 FT REPEAT 155 180  
 FT REPEAT 181 206  
 FT REPEAT 207 232  
 FT REPEAT 233 258  
 FT REPEAT 259 284  
 FT REPEAT 285 310  
 FT REPEAT 311 336  
 FT REPEAT 365 391  
 FT CONFLICT 62 62 I -> T (IN REF. 4).  
 FT CONFLICT 77 77 G -> V (IN REF. 1).  
 FT CONFLICT 81 81 R -> K (IN REF. 4).  
 FT CONFLICT 174 174 I -> W (IN REF. 5).  
 FT CONFLICT 320 320 S -> P (IN REF. 6).  
 SQ SEQUENCE 423 AA; 47073 MW; BEGF82438512.B4E CRC64;  
 Query Match 5.4%; Score 170.5; DB 1; Length 423;  
 Best Local Similarity 23.4%; Pred. No. 1.1e-05;  
 Matches 85; Conservative 54; Mismatches 136; Indels 89; Gaps 15;  
 126 LKSLHFRMI-VKDSLDQNLARDGRHVLHAKLKDCKSGFTTDGLFHIGRFCKSLRVLFLE 184

Db 80 LRKLSRGICIGVGDSSIKTFAQNCRIEH-LNLANGCTKTIIDTCYSLSRFCSKXKLHLDLT 138  
 Qy 185 ES-SILEKGEWHEALANNVTLETNFYITDIADVVKIEDLELLAKNCPNLVSVKLTDC- 242  
 Db 139 SCVSIITSSLSKGISGECGRN---LEYLNLSMCD--QITKDGIEALVRCRGKALLRGCT 193  
 Qy 243 ----EILDVNFKHA-----SAAEEFCGGTYNEEPERYSAISLP--AKLC 282  
 Db 194 QLEDEALKHIONQYCHELVSLNLSQCSRITDEGVVQICRGCH-----RLQALCLSGCSNLT 248  
 Qy 283 RLGLTITIGKNELPIVFMFAAVLKKLDLLYAMLTDEHMLIORCPNLEVLTRNV--IGD 340  
 Db 249 DASITALGLN-----CPRLQILLEAARCSHLTD 275  
 Qy 341 RLGEVLGRCCRLKRLRIERIGDDDDQMEDEGVTVSHRGLIALSGCCSELEYMAV-YVSDI 399  
 Db 276 AGFTLLARNCHELEKMDLE-----ECILITDSTLIQLSHCPKQLQALSJHCELI 325  
 Qy 400 TNASLEHIGTHLKNLDCDFRLVLDHEBKITDLPDNGVRALLRGCDKLRFPALY----LR 455  
 Db 326 TDGDIHLNSNTOGHERLRLVLELDNCLLITDVALEH-----LENCRGRLRLDYDCCQVT 380  
 Qy 456 RGGL 459  
 Db 381 RAGI 384

RESULT 7  
 ID\_POF2 SCHPO STANDARD; PRT; 463 AA.  
 AC 074783;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE F-box/LRR-repeat protein 2 (F-box and leucine-rich repeat protein 2)  
 GN (F-box protein pof2).  
 GN POF2 OR SPBC25B2.11.  
 OS Schizosaccharomyces pombe (Fission Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MEDLINE=21668955; PubMed=11809834;  
 RA Katayama S., Kitamura K., Lehmann A., Nikaido O., Toda T.;  
 RT "Fission Yeast F-box protein Pof3 is required for genome integrity and  
 RT telomere function.";  
 RL Mol. Biol. Cell 13:211-224 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jørgensen K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstaere E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB062778; BAB50689.1; -;  
 DR EMBL; AL031863; CAA21269.1; -;  
 DR PIR; T39987; T39987.  
 DR GeneDB SPombe; SPBC25B2.11; -;  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR007089; LRR cys.  
 DR InterPro; IPR006553; LRR cys sub.  
 DR InterPro; IPR008945; Skp1\_Skp2.  
 DR Pfam; PF00560; LRR; 1.  
 DR SMART; SM00367; LRR; CC; 1.  
 DR PROSITE; PS50181; FBOX; 1.  
 KW Ub1 conjugation pathway; Repeat; Leucine-rich repeat.  
 FT DOMAIN 1 42 F-BOX.  
 FT REPEAT 93 118 LRR 1.  
 FT REPEAT 145 170 LRR 2.  
 FT REPEAT 171 196 LRR 3.  
 FT REPEAT 276 301 LRR 4.  
 FT REPEAT 302 327 LRR 5.  
 FT REPEAT 328 353 LRR 6.  
 FT REPEAT 354 378 LRR 7.  
 FT REPEAT 380 405 LRR 8.  
 SQ SEQUENCE 463 AA; 52360 MW; 07435B3C2BBB2073 CRC64;  
 Query Match 4.8%; Score 154; DB 1; Length 463;  
 Best Local Similarity 20.5%; Pred. NO. 0.00024;  
 Matches 97; Conservative 70; Mismatches 145; Indels 162; Gaps 22;  
 QY 86 HLESILKLGKPRAMFNIPEDWGHHVTP--VWK-----EISQYFCLK-----SLHF 131  
 DB 14 VLEADELRCK-----STVCTSRNFIPTLWEKVVFQNEAQLNFFDTLQYSKDVSYF 67  
 QY 132 RMIVKDSQLQNLARDRGVHLAKDKSGFTTGLFHIGRFCKSLRVLFLEESSILEK 191  
 DB 68 R-----YLFKLCNSVRKFLTDK--HLMLMTLATGTSRLNLSGCTRI 107  
 QY 192 DGEWLHEALNNTVLETNLYTIDTAVVKIELELLAKNCPLVSKLTDCCILDLVNF 251  
 DB 108 SEPLTGKLYQNLNLTINF--SNIFSPANILEYISDNCPEKALNIGCNGLVE----- 160  
 QY 252 KHASALEEFCGCTYNEEPERYSAISLPAPKLRLGLTYIGKNELPIVFVFAAFLKKLDL 311  
 DB 161 -----EKEDLFDLIS 205  
 QY 312 AMLDTEHCHLQPCNPVLETRNV--IGDRGLEVLGRCKRLKRLIERGD-----DDQ 365  
 DB 366 GME--DEECTVS-----HRLIALS--QGCELEYMAVYVDITNASLEHI--GTHLKNLKD 416  
 DB 206 GEGEFNADTLRLVSRNGLKELSDMGCTELSHITFTFN--LNCELDAMRALSLNNLPD 263  
 QY 417 FRLVLIDHEKITDLPDNGVRALLRGCDLRRFALYLRFG--GLTDVGLGYIGQYSPNVR 475  
 DB 264 LK-----DSDIELITCKFSKLN--SLFLSKCIGLTDSSLSLSTKLQSLOT 306

QY 476 WMLLGYGVE-SDAGLLEFAKGPSLQKLEMEGCLFFSERALAVAAATOLTSRLYL 528  
 DB 307 TLHGHGVEITDGVQCILKSKRNITYIDFGCLRLSD-----IAVSAIAKLPYL 356  
 RESULT 8  
 FL2L RAT  
 ID FL2L RAT STANDARD; PRT; 276 AA.  
 AC Q9QZH7;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE F-box/LRR-repeat protein 2-like.  
 GN FBL2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., INTERACTION WITH SKP1, SUBCELLULAR LOCATION, AND  
 RP TISSUE SPECIFICITY.  
 RC STRAIN=Buffalo; TISSUE=Liver;  
 RX MEDLINE=99439713; PubMed=10508920;  
 RA "Ilvin G.P., Riialand M., Glaiese D., Guguen-Guillouzo C.;  
 RT "Identification of a novel Skp2-like mammalian protein containing  
 RT F-box and leucine-rich repeats.";  
 RL FEBS Lett. 459:75-79(1999).  
 CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated  
 CC proteins and promotes their ubiquitination and degradation.  
 CC -!- SUBUNIT: Interacts with SKP1. Part of a SCF (SKP1-cullin-F-box)  
 CC protein ligase complex (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- TISSUE SPECIFICITY: Widely expressed, with highest expression in  
 CC skeletal muscle, heart and brain.  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF182443; AAF01221.1; -;  
 DR PIR; T52349; T52349.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR007089; LRR cys.  
 DR InterPro; IPR006553; LRR\_cys\_sub.  
 DR Pfam; PF00646; F-box; 1.  
 DR SMART; SM00256; FBOX; 1.  
 DR SMART; SM00367; LRR; CC; 1.  
 DR PROSITE; PS50181; FBOX; 1.  
 KW Ub1 conjugation pathway; Leucine-rich repeat; Repeat.  
 FT DOMAIN 22 68 F-BOX.  
 FT REPEAT 90 115 LRR 1.  
 FT REPEAT 142 167 LRR 2.  
 FT REPEAT 168 193 LRR 3.  
 FT REPEAT 194 219 LRR 4.  
 FT REPEAT 220 245 LRR 5.  
 FT REPEAT 246 271 LRR 6.  
 SQ SEQUENCE 276 AA; 30460 MW; 982994F91A265E1E CRC64;  
 Query Match 4.6%; Score 147.5; DB 1; Length 276;  
 Best Local Similarity 23.0%; Pred. NO. 0.00037;  
 Matches 70; Conservative 42; Mismatches 98; Indels 95; Gaps 13;  
 QY 233 NLVSVKLTDCILDLVNFVKASA-----LSEFCGCTYNEEPERYSAISLPAPKLRLGLT 287  
 DB 56 NVLADGSGNWQRIDLDFQDIEGRVVENISKRCGGFLR-----KLSIRGLCL 102

QY 288 YIGKNEPVPVMPFAVVKLLDLYAMLDTEHCHMLIQPCNLEVLTRNV--IGDRGLEV 345  
 Db 103 GVGGNALR---TFA-----QNCRNIEVLSUNGCTKTDTACTS 137  
 QY 346 LGRCKKIKRLIRERGDGQMEDECTVSHRGLIALSQSCSELEYMAVYVDITNASLE 405  
 Db 138 LSKFCSKLRHLDLASC-----SITWMSLKALSEGCPLEQLNI----- 176  
 QY 406 HIGTHLKNCDPRLVLLDHEKIDTLPDNGVRLLRGCDKLRFAVLRGGIT---DV 462  
 Db 177 -----SWCD-----QVT-----KGIQALVRGCGGLK--ALFLK--GCTOLEDE 211  
 QY 463 GLGVIGQYSPN-VRMMLLYGVSAGLEFAKCPISQKLEMGCLFFFSERALAAVATQ 521  
 Db 212 ALKYIGAHCPBLVTNLQTCITQITDEGLTTCRCGHKLQSLCAGSCNITDALNALGON 271  
 QY 522 LLSLR 526  
 Db 272 CPLRLR 276

RESULT 9  
 ID\_FXL6 HUMAN STANDARD; PRT; 539 AA.  
 AC Q8NS31; Q8NS49; Q9UKC7;  
 DT 10-OCT-2003 (Rel. 42, last sequence update)  
 DT 15-MAR-2004 (Rel. 43, last annotation update)  
 DE F-box/LRR-repeat protein 6 (F-box and leucine-rich repeat protein 6).  
 DE FBXL6 OR FBL6.  
 GN FBXL6 (Human)  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE-Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., Suzuki Y., Oiyashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 90-372 FROM N.A. (ISOFORM 1/2).  
 RX MEDLINE=20030360; PubMed=10531035;  
 RA Cenciarelli C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M., Pagano M.;  
 RT "Identification of a family of human F-box proteins."  
 RL Curr. Biol. 9:1177-1179 (1999).

CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated proteins and promotes their ubiquitination and degradation.  
 CC -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex (by similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q8NS31-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8NS31-2; Sequence=VSP\_008498;  
 CC Note=No experimental evidence available;  
 CC -!- SIMILARITY: Contains 1 F-box domain.  
 CC -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; BC033066; AAH33066.1; -;  
 CC EMBL; AK026541; BABIS499.1; -;  
 CC EMBL; AF174592; AAF04513.1; -;  
 CC Genbank; HGNC:13603; FBXL6.  
 CC InterPro; IPR001810; F-box.  
 CC InterPro; IPR001611; LRR.  
 CC InterPro; IPR006553; LRR\_Cys\_sub.  
 CC InterPro; IPR008945; Skp1\_Skp2.  
 CC Pfam; PF00646; F-box; 1.  
 CC Pfam; PF00560; LRR; 1.  
 CC SMART; SM00367; LRR\_CC; 3.  
 CC PROSITE; PS0181; PROX; FALSE NEG.  
 CC Ubl conjugation pathway; Leucine-rich repeat; Repeat;  
 KW Alternative splicing.  
 FT DOMAIN 10 162 F-BOX.  
 FT REPEAT 219 244 LRR 1.  
 FT REPEAT 383 408 LRR 2.  
 FT REPEAT 497 528 LRR 3.  
 FT VARSPPLIC 208 213 Missing (in isoform 2).  
 FT /FTid=VSP\_008498.  
 SQ SEQUENCE 539 AA; 58587 MW; 815318544653BB2 CRC64;  
 Query Match 4.4%; Score 141; DB 1; Length 539;  
 Best Local Similarity 21.1%; Pred. NO. 0.0029;  
 Matches 118; Conservative 62; Mismatches 162; Indels 218; Gaps 23;  
 QY 31 VVLDGVIPYIDDPKORDAVSQVCRWYELDSLTRKHVTIALCYTTTPARLRRRFFHLESL 90  
 Db 128 VAADGPMFPL-----GRAARVCRWQEAASQP-----ALWHTVTLSPVGRPAKGV 175  
 QY 91 KLKGGPRAAMFNLIPE-----DWGGHVTVPWKEISQYFDCLKSLHFRMIVKDS 139  
 Db 176 KAEKKLLASLEWMPNRSQQLRLTLHWKSVQHPVLKVG---ECCPLTF----- 224  
 QY 140 DLQNLARDRGVHLHALDKKSGFTTGLFHIGRCCKSLRVFLRES-----SILEK 191  
 Db 225 -----LKLGGCHGVTDALVLMKAKCCQLHSLDLQHSVMVSTAVVSFLEE 269  
 QY 192 DGEWHELALNNTVLETLNLYTDIAVVKIEDELLAKNCPNLVSKLTDCLDLVNPFF 251  
 Db 270 AGSRMRKWLTVSSQT-----AALG-----ALLGCCFPL-----QVLEV----- 305  
 QY 252 KHASALEEFCGGTYNEEPERYSAISLPKACLRLGTYIGKNELPIVFMFAVKKLDLY 311  
 Db 306 -----STGINRNSIPLQLPVEALQK----- 325  
 QY 312 AMLQTEDHCHMLIQPCNLEVLTRNVIGDRGLEVLRGCCCKRLKRLRIERGDDDDQMEDEE 371  
 Db 326 -----GCPQLQVLRLLNL----- 345  
 QY 372 GTVSHRGLIALSQSCSELEYMAVYVDITNASLEHIGTHLKNCDPRLVLLDHEKIDTDL 431

```

Db 346 G-----RG-VAPGPGFSPLEBELCLASSTCNFVSNEVLGRLLHSGPNLRL----- 388
QY 432 PLNGVALLRGCDKLRFAFYLRGGGLTDVGLGYIGQYSPNVWMLLYGVGSDAGLLE 491
Db 389 -LD-----LRGCR-----ITPAGLDLPQRELEQ-----LHLGYGTSDBLTLA 427
QY 492 -----FAKGCPSLOKMRGCLFFSERAL--AVAAQTUT-----SLRYLWVQYGVGS 536
Db 428 KESGPFUTQWCHTLRELDLSG-QGFSKDLQALAAFLSTPGGSHPALCSNLNRTV 486
QY 537 PSGRDLLVWARP---FWNIE 553
Db 487 PSTVSSVISSCPGLLYNLLE 506

RESULT 10
ESAS TRYEQ STANDARD; PRT; 630 AA.
AC P2637;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative adenylate cyclase regulatory protein.
GN ESAG8C.
OS Trypanosoma equiperdum.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5694;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL;
RX MEDLINE=91293076; PubMed=2065652;
RA Ross D.T., Raibaud A., Florent I.C., Sather S., Gross M.K.,
RT "The trypanosome VSG expression site encodes adenylate cyclase and a
RL leucine-rich putative regulatory gene.";
EMBO J. 10:2047-2053 (1991).
CC -!- FUNCTION: May interact with adenylate cyclase to regulate its
CC activity.
CC -!- FUNCTION: May be involved in the posttranscriptional regulation of
CC genes in VSG expression sites.
CC -!- DEVELOPMENTAL STAGE: Expressed only in the bloodstream form of the
CC parasite.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59385; CAA42028.1; -.
CC PIR; S16358; BWU78Q.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00560; LRR; 9.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Repeat; Leucine-rich repeat; cAMP biosynthesis; Zinc-finger;
KW DNA-binding.
KW ZN_FING 10 46 RING-TYPE.
FT DOMAIN 47 54 ARG/LYS-RICH (BASIC).
FT DOMAIN 71 77 ARG/LYS-RICH (BASIC).
FT REPEAT 108 153 LRR 1.
FT REPEAT 181 225 LRR 2.
FT REPEAT 252 296 LRR 3.
FT REPEAT 322 366 LRR 4.
FT REPEAT 368 412 LRR 5.

```

```

FT REPEAT 414 458 LRR 6.
FT REPEAT 460 504 LRR 7.
FT REPEAT 506 550 LRR 8.
FT REPEAT 552 595 LRR 9.
SQ SEQUENCE 630 AA; 69989 MW; F882E49A6A6E6651 CRC64;

Query Match 4.4%; Score 139; DB 1; Length 630;
Best Local Similarity 20.9%; Pred. No. 0.0052;
Matches 110; Conservative 67; Mismatches 157; Indels 192; Gaps 29;

QY 126 LKSLHFRMIVKD-----SDLQNLARDGRGHVHLKLDKCSG 162
Db 161 LRKLRMKRTMVNDMWCSSISGLLKPLVHLEVDGGRGVTDTGLCRK--TLEALSDESCIN 218
QY 163 FTDTGLFHIGRFC--KSLRVLFLEESSILEKGEWLHEALNNVTLETNF----YLTDI 216
Db 219 ITKG-----FDKICALPQLTSLSCQTNVTDKLRCHP-----DGKLVLAYSSCHETDL 270
QY 217 AVV-----KIEDLEL-----LAKNC--PNLVSVKLTDCLEL-----DLVNF----- 250
Db 271 TAICGMRSLKLSLSCGWNVTKGLLECKFSNRLDLSGLVLGSAVLKNLNLKVL 330
QY 251 -----FKHASALEEF-----CGGTYNEPERYSALSIPAKLCRLGLTYIGKNEL 294
Db 331 VSNCKNFKDLNGLERLVNLDKLNLSGCHGV-----SSLGFWANLSNLKELDISCES 382
QY 295 PIVFMFAAVLKLDLYAMLDTEHCHMLIQRCNPLEVLETRNVIGDRGLEVLGRCCRLK 354
Db 383 LVCFDGLQDLNNLEVLY-LRDVKSF-----TNVGAIKNLSKMRLELDLSG--CERIT 430
QY 355 RLRIERGGDDQGMEDDEGTVSHRGLIALS-QGCSLEL-----EYMAVYVSDITN-- 401
Db 431 SL-----SGLETLKGLELSLEGCGEIMSPDPIWSLHLRLVLYSECGNLE 476
QY 402 -----ASLEHIGTH-----LKNLCDFRLVLDHHEKITDPLDNGVRAL 440
Db 477 DLGLEGITGLEELYHLHGCKRCKTNFGPIWNLNRNVCVVELSCENLELDLSGLQCLTGLEEL 536
QY 441 -LRGCDKLRRFALYLRGGGLTDVGLGYIGQYSPNV-----WMLLYGVGSDAGLLE 491
Db 537 YLIGCEI-----TPIGVVG-----NLRNLKCLSTCWC-----ANLKE 569
QY 492 FA--KGCPSPQKLEMRGCLFFSERALAAVATLTSLRYL-WVQYGV 534
Db 570 LGGDLRLVNLKLDLSGCCGLSSVF---MELMSLPKQLQWIFYGFG 611

RESULT 11
ESAS TRYBB STANDARD; PRT; 630 AA.
AC P23799;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative adenylate cyclase regulatory protein (Leucine repeat protein)
DE (VSG expression site-associated protein Fl4.9).
GN ESAG8 OR T-LR.
OS Trypanosoma brucei brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EATRO 164;
RX MEDLINE=91061751; PubMed=2247064;
RA Smiley B.L., Stadnyk A.W., Myler P.J., Stuart K.;
RT "The trypanosome leucine repeat gene in the variant surface
RT glycoprotein expression site encodes a putative metal-binding domain
RT and a region resembling protein-binding domains of yeast, Drosophila,
RL and mammalian proteins.";
RN Mol. Cell. Biol. 10:6436-6444 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=EATRO 1125;

```

RA MEDLINE=91081317; PubMed=2259625;  
RA Lips S., Revelard P., Pays E.;  
RT "A gene from the VSG expression site of Trypanosoma brucei encodes a  
RT protein with both leucine-rich repeats and a putative zinc finger.";  
RL Nucleic Acids Res. 18:7299-7303(1990).  
CC -!- FUNCTION: May interact with adenylate cyclase to regulate its  
CC activity.  
CC  
CC -!- FUNCTION: May be involved in the posttranscriptional regulation of  
CC genes in VSG expression sites.  
CC  
CC -!- DEVELOPMENTAL STAGE: Expressed only in the bloodstream form  
CC of the parasite.  
CC  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
CC  
CC -!- SIMILARITY: Contains 23 leucine-rich (LRR) repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; M58701; AAA32117.2; ALT\_TERM.  
CC  
CC EMBL; X55978; CRA39448.1; -  
CC PIR; A36359; A36359.  
CC PIR; S13724; S13724.  
CC InterPro; IPR001611; LRR.  
CC InterPro; IPR001841; Znf\_ring.  
CC Pfam; PF00560; LRR; 8.  
CC Pfam; PF00097; zf-C3HC4; 1.  
CC SMART; SM00184; RING; 1.  
CC PROSITE; PS00518; ZF\_RING\_1; 1.  
CC PROSITE; PS00089; ZF\_RING\_2; 1.  
CC Repeat; Leucine-rich repeat; CAMP biosynthesis; Zinc-finger;  
KW DNA-binding.  
FT ZN FING 10 46 RING-TYPE.  
FT DONAIN 47 54 ARG/LYS-RICH (BASIC).  
FT DONAIN 71 77 ARG/LYS-RICH (BASIC).  
FT REPEAT 115 123 LRR 1.  
FT REPEAT 124 148 LRR 2.  
FT REPEAT 149 170 LRR 3.  
FT REPEAT 171 195 LRR 4.  
FT REPEAT 196 218 LRR 5.  
FT REPEAT 219 242 LRR 6.  
FT REPEAT 243 266 LRR 7.  
FT REPEAT 267 289 LRR 8.  
FT REPEAT 290 313 LRR 9.  
FT REPEAT 314 336 LRR 10.  
FT REPEAT 337 359 LRR 11.  
FT REPEAT 360 382 LRR 12.  
FT REPEAT 383 405 LRR 13.  
FT REPEAT 406 428 LRR 14.  
FT REPEAT 429 451 LRR 15.  
FT REPEAT 452 474 LRR 16.  
FT REPEAT 475 497 LRR 17.  
FT REPEAT 498 520 LRR 18.  
FT REPEAT 521 543 LRR 19.  
FT REPEAT 544 566 LRR 20.  
FT REPEAT 567 589 LRR 21.  
FT REPEAT 590 613 LRR 22.  
FT REPEAT 614 630 LRR 23.  
FT CONFLICT 22 22 V -> L (IN REF. 2).  
FT CONFLICT 25 25 L -> E (IN REF. 2).  
FT CONFLICT 38 38 Q -> E (IN REF. 2).  
FT CONFLICT 203 203 F -> C (IN REF. 2).  
FT CONFLICT 215 215 N -> S (IN REF. 2).  
FT CONFLICT 258 258 M -> V (IN REF. 2).  
FT CONFLICT 312 312 L -> P (IN REF. 2).  
FT CONFLICT 321 321 K -> R (IN REF. 2).  
FT CONFLICT 321 321 T -> S (IN REF. 2).  
FT CONFLICT 437 437 Y -> H (IN REF. 2).  
FT CONFLICT 462 462 Y -> L (IN REF. 2).  
FT CONFLICT 490 490 M -> L (IN REF. 2).  
FT CONFLICT 504 504 IWN -> FGI (IN REF. 2).

FT CONFLICT 511 511 C -> L (IN REF. 2).  
FT CONFLICT 522 522 D -> E (IN REF. 2).  
FT CONFLICT 543 546 EITT -> KIOP (IN REF. 2).  
FT CONFLICT 549 549 V -> I (IN REF. 2).  
FT CONFLICT 582 582 L -> V (IN REF. 2).  
FT CONFLICT 620 620 K -> E (IN REF. 2).  
SQ SEQUENCE 630 AA; 69998 MW; A65A35B5DCB50F7E CRC64;  
  
Query Match 4.2%; Score 133; DB 1; Length 630;  
Best Local Similarity 20.7%; Pred.No. 0.015;  
Matches 109; Conservative 69; Mismatches 156; Indels 192; Gaps 29;  
  
Qy 126 LKSLHFRMIVKD-----SDQLNLPDRGHVHALKDKCSG 162  
Db 161 LRKLKMKRTWVNDWMCSSIGLLKFLVHLEVDGSRGVTDTGLFKL--TLEALSLDNCIN 218  
Qy 163 FTTDGLPHIGRFC--KSLRVLFLEESILEKDEGLHMLANNVLETFNF---YLTDI 216  
Db 219 ITK3---PDKICAPQLTSLSLCQTNVTKDLRCHHP---DGKLMLODISSCHEITDL 270  
Qy 217 AVV---KIDLEL-----LAKNC--PNLVSVKLTDCIIL-----DLVNF--- 250  
Db 271 TAIGGVRSLEKLSLSCGWNVTGKLEELCKFSNARELDISGCLVLSAVLKNLNLKVL 330  
Qy 251 -----FKHASALEEF-----CGTYNEPPEPYSALSIPAKLCRLGLTYIGKNEL 294  
Db 331 VSNCKNFQDLNGLERLVNLEKLNLSGCHGV-----SSLGFVANLSNLKELDISGCES 382  
Qy 295 PIVMFAAVLKLDLILYANLDTEDHCLMIQPCPNLEVTENVIGDRGLEVLGRCCRELK 354  
Db 383 LVCPDGLQDLNLEVLV-LRDVKSF-----TNVGAIKNLSKMRLELDLSG--CERIT 430  
Qy 355 RLRIERGDDQGMEDBEGTVSHRGLIALS--OGCSEL-----EYMAVYVSDITN-- 401  
Db 431 SL-----SGLETIKGLELSLEGCGEIMSPDPIWSLYHLRVLYVSECGNLE 476  
Qy 402 -----ASLEHIGTH-----LKNICDFRLVLLDHEEKITDPLDNGVRAL 440  
Db 477 DLSGLQCLTGLEEMVYLGCRKCTNFGPIWNLNRVLCVLELSCCNLDLDSGLQCLTGLEEL 536  
Qy 441 -LRGCDKLRRFALYLRGGLTDVGLGYIGQVSPNVV-----WMLLYGVGESDAGLLE 491  
Db 537 YLIGCEEI-----TTIGVVG--NLRNLKCLSTCWC-----ANLKE 569  
Qy 492 FA--KGCPSLQKLEMRGCLFESERALAVAATQLTSRLYL--WVOGYG 534  
Db 570 LGGHLRLVNLKLDLSGCCGLSSSVF-----MELMSLPKLQWIFYGFG 611  
  
RESULT 12  
RAD7 YEAST  
ID RAD7 YEAST STANDARD; PRT; 565 AA.  
AC P06779;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE DNA repair protein RAD7.  
GN RAD7 OR YJR052W OR J1665.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP MEDLINE=90236305; PubMed=2158927;  
RA Melnick L., Sherman F.;  
RT "Nucleotide sequence of the COR region: a cluster of six genes in the  
RL yeast Saccharomyces cerevisiae.";  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87064430; PubMed=3023893;  
RA Perozzi G., Prakash S.;



"RAD7 gene of Saccharomyces cerevisiae: transcripts, nucleotide sequence analysis, and functional relationship between the RAD7 and RAD23 gene products." Mol. Cell. Biol. 6:1497-1507 (1986).

[3]  
SEQUENCE FROM N.A.  
STRAIN=S288C;  
Huang M.-E., Manus V., Chuat J.-C., Galibert F.;  
"Revised nucleotide sequence of the COR region of yeast Saccharomyces cerevisiae chromosome X." Yeast 10:811-818 (1994).

CC -!- FUNCTION: This protein is one of 10 proteins (RAD1, 2, 3, 4, 7, 10, 14, 16, 23 and MMS19) involved in excision repair of DNA damaged with UV light, bulky adducts, or cross-linking agents.

CC -!- MISCELLANEOUS: Mutants with mutations in the RAD7, RAD14, RAD16, and RAD23 genes show partial incision defectiveness.

CC -!- SIMILARITY: TO S.POMBE SPC613.14.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; M37696; AAB59347.1; -  
EMBL; M13015; AAB34953.1; -  
EMBL; L26347; AAB62860.1; -  
EMBL; L36344; AAB88755.1; -  
EMBL; L249552; CAA89580.1; -  
PIR; A25226; A25226.  
GeneOnline; 141885; -  
SGP; S0003813; RAD7.  
DR GO; GO:0000113; C:nucleotide excision repair factor 4 complex; IDA.  
DR GO; GO:0000108; C:repairosome; IDA.  
DR GO; GO:0003677; F:DNA binding; IDA.  
DR GO; GO:0008094; F:DNA dependent ATPase activity; IDA.  
DR InterPro; IPR007089; LRR\_cys.  
KW DNA repair.  
DR DOMAIN 1 200 HYDROPHILIC.  
FT CONFLICT 278 279 LL -> FV (IN REF. 1).  
FT CONFLICT 504 505 AC -> RP (IN REF. 1).  
FT SEQUENCE 565 AA; 63777 MW; 85C77357DC99737A CRC64;  
Query Match 4.2%; Score 132; DB 1; Length 565;  
Best Local Similarity 23.1%; Pred No. 0.016;  
Matches 114; Conservative 78; Mismatches 188; Indels 114; Gaps 24;  
60 DSLTKKHVTIALCYTTTPARLRERPHLESKLKGPRAAMNLPEDWGGHVTWVKEI 119  
131 ESLTKRQNTAKIQN---RRRKRAADLLDRVRNVKVSLSQSLCITTKISENISKWQEA 187  
120 SQYFCLKSLHFRMI-----VKSDLQNLA---RDRGHVLHAKLKDCKSGFTTDLGFH 170  
188 ---DESSKLVNKLRLDVLGGVSTANLNNLAKALSKNALNDHTLQL----- 230  
171 IGRFCKS-LRVLFFLESSILEKDGWELHNLNNTVLTNPLYTDIAV-----VKIEDL 224  
231 ---FLKTDLKRUTTFSDCKISFDG-----YKTLAIFSHLTSLQMCQGLNHSL 278  
225 ELLAKNCPNLVSKLTDCEILD---LVN-----FFKHASALEFPCGGTYNEEPVSA 274  
279 LVIAEKLPLNKSIN-----LDGPFLLNEDTWKFPVINKGRLEEF---HISNTHRFTD 328  
275 ISLPAKLKRLGTYIGKNELPVPFAVLKLDLL--YAML-----DTEDHCLMIIQPCP 327  
329 KSLSNLLNCGSTLVSLG-----LSRLDSISNVALLPQVLVDNDFHSLCLIEYFP 377  
328 NLEVLTRNVIGDGLVIGRCCKRLKRLRIERGDDDDQGMDEEGTVSHRGLIALSQGCS 387  
378 NEEDVNDIITII-----NLIGQIGRTURKLV-----NGCIDLTDNMIINGLTAFIPEKC 426

QY 388 ELEYMAVYVSD-ITNASLREHI--GTHLKNL--CDFRLVLIDHEEKITDLPDNGVRALLR 442  
Db 427 PLEVLSLEESDQITTSLSLSTFFSKVELNNLIESFRCLQGLDWAIIELLL-NGARDSLR 485  
QY 443 GCDKLRFRFALYLRGGGLTDVGL-GYIGQVSPNVNRMMLLVGVGSDAGLLS-FAKGCPISLQ 500  
Db 486 S-----LNLSLKEITKEAFVALACPNLTLDLGFVRCVDDSVIQLMGQENPFLT 535  
QY 501 KLEMRGCLFFSERA 514  
Db 536 VIDVFGDNLVTEKA 549

RESULT 13  
NAL2 HUMAN STANDARD; PRT; 1061 AA.  
AC P59046;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE NACHI-, LRR- and PYD-containing protein 12 (PYRIN-containing APAF1-like protein 7) (Monarch-1).  
GN NALP12 OR PYPAF7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RP MEDLINE=22162427; PubMed=12019269;  
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S., Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;  
"PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates activation of NF-kappa B and caspase-1-dependent cytokine processing." J. Biol. Chem. 277:29874-29880 (2002).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=22451042; PubMed=12563287;  
RA Teschopp J., Martinon F., Burns K.;  
"NALPs: a novel protein family involved in inflammation." Nat. Rev. Mol. Cell Biol. 4:95-104 (2003).  
RN [3]  
RN SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RC TISSUE=Lymphoma;  
RA Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;  
"Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases."  
RN [4]  
RN SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Leukocyte;  
RA MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millatny S.J., Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grinchwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RT -!- FUNCTION: May mediate activation of CASP1 via ASC and promote



induces glucose repression. Is not an essential protein. Involved in substrate recognition in ubiquitin-dependent degradation.

-!- PATHWAY: Necessary for the glucose repression pathway.

-!- SUBUNIT: Part of a SCF E3 ubiquitin ligase complex containing CBP3, CDC53, HRT1 and GRI1.

-!- SUBCELLULAR LOCATION: Associated with the particulate fraction. Probably forms a complex by protein-protein interactions via its leucine-rich segment.

-!- INDUCTION: Expressed constitutively at low levels.

-!- SIMILARITY: Contains 1 F-box domain.

-!- SIMILARITY: Contains 12 leucine-rich (LRR) repeats.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; M59247; AAA34652.1; -  
 EMBL; Z49590; CAA89617.1; -  
 EMBL; L47993; AAB39313.1; -  
 FTR; A41529; A41529.  
 GenOnline; 141922; -  
 SGD; S0003850; GRI1.  
 GO; GO:0000142; C:contractile ring (sensu Saccharomycetes); IDA.  
 GO; GO:0005737; C:cytoplasm; IDA.  
 GO; GO:0005634; C:nucleus; IDA.  
 InterPro; IPR001810; F-box.  
 InterPro; IPR001611; LRR.  
 InterPro; IPR007089; LRR\_cys.  
 Pfam; PF00646; F-box; 1.  
 Pfam; PF00560; LRR; 1.  
 SMART; SM00256; FBOX; 1.  
 PROSITE; PS0181; FBOX; 1.  
 Glucose metabolism; Ubl conjugation pathway; Repeat;  
 Leucine-rich repeat.  
 DOMAIN 314 361 F-BOX.  
 LRR 1. 434  
 LRR 2. 435  
 LRR 3. 460  
 LRR 4. 486  
 LRR 5. 512  
 LRR 6. 548  
 LRR 7. 567  
 LRR 8. 593  
 LRR 9. 619  
 LRR 10. 645  
 LRR 11. 670  
 LRR 12. 726  
 POLY-ASN.  
 ASN-RICH.  
 DOMAIN 1045 1124  
 SEQUENCE 1151 AA; 132733 MW; 6BB6C46611B6F825 CRC64;

Query Match 4.0%; Score 127.5; DB 1; Length 1151;  
 Best Local Similarity 20.7%; Pred. No. 0.091;  
 Matches 119; Conservative 78; Mismatches 190; Indels 189; Gaps 29;

QY 11 LSLRSMTEERNVTRVVDVLDVIPPY-----IDDPKRDV----- 49  
 DB 212 MRLRALETENMEIRNLRL--KILTIIEEYKSLYAYCHSKLRGQQQVENPTDNFIWINSI 269  
 QY 50 -----SQVCRWVELDSLTKRHVTIAL-----CYTTTPARLRRFP----- 85  
 DB 270 DTTESDLKGLQDLKRYGRFINNVLSNPSNQICTSVT-----RRSPFALNMLPSII 324  
 QY 86 -HLESKLKXGKPRAMFNILPEPDGCVTPWVKEISQYFDCLKSLHFRMIVKDSLDQNL 144  
 DB 325 LHLILDKLQKQYDIVKFLTVSKLWA-----EI-----IVKILYRPHINKSKSLDLF 371  
 QY 145 ARDRGHVHALKDKCSGGFTDGLFHIGRFCKSLRVLFLEESSILEKDGHWLHNLNT 204

372 -----LRMKLT-----SEETVFNRLMKRLNFSV-----GDYMHDTLNIF 410  
 QY 205 VLETLNLYTDIAVVKIEDLELLAKNCPNLVSKLTDCEILDVNFPHASALEEFCGGT 264  
 DB 411 V-----GCKNL-----ERLTLV-FCKHTSV----- 430  
 QY 265 YNEPERYSATSLPAKLCRLGLTVIGKNELPIVFMFAAVLKLDLKYAMLDTEHCM--L 322  
 DB 431 -----PISAVLRCKKF-----LQSVDT-GIRDVSDVDFTL 461  
 QY 323 IQRCPNLE-----VLETRNVIGD--RGLEVLGRCKRLKRLRIERGGDDQGHDESGTVSHR 377  
 DB 462 ATVCPRVGVFPQARNVTFDSLRFVHS---PMLAKRIKITANN--MNDE----- 508  
 QY 378 GLI-ALSOQSELEYMAVYVS-DITNASLEHGHAKNLCDFRVLVLDHBEKITDPLDN 435  
 DB 509 -LVELLANKCPLLVEVDITLSPNVTSSLLKLRLVQLREFRIT---HNTNIT-----DN 560  
 QY 436 GVRALLRGCDKLRRFPALYLRG--GLTDVGIGYIGQYSPNVRMMLGYVGE-SDAGLEF 492  
 DB 561 LFQELSKVDDMPSLRLLDLSGCENITDKTIESIVNLAPKLRNVFLGKCSRITDASLFQL 620  
 QY 493 AKGCPSLQKLEMRGCLFPPSERALAAVATQTLSTLYL 528  
 DB 621 SKLGKNLQTVHFGHCFNITDNGVRLFHSCTRIQYV 656

RESULT 15  
 FXL4 HUMAN  
 ID FXL4 HUMAN STANDARD; PRT; 621 AA.  
 AC Q9UKA2; O95919; Q9UJU0.  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE F-box/LRR-repeat protein 4 (F-box and leucine-rich repeat protein 4)  
 DE (F-box protein FBL5).  
 GN FBL4 OR FBL4 OR FBL5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=2003061; PubMed=10531037;  
 RA Winston J.T., Koepf D.M., Zhu C., Elledge S.J., Harper J.W.;  
 RT "A family of mammalian F-box proteins."  
 RL Curr. Biol. 9:1180-1182 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2003060; PubMed=10531035;  
 RA Cenciarelli C., Chiar D.S., Guardavaccaro D., Parks W., Vidal M.,  
 RA Pagano M.;  
 RT "Identification of a family of human F-box proteins."  
 RL Curr. Biol. 9:1177-1179 (1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20399565; PubMed=10945468;  
 RA Ilyin G.P., Riialand M., Pigeon C., Gugen-Guillouzo C.;  
 RT "cDNA cloning and expression analysis of new members of the mammalian  
 RT F-box protein family."  
 RL Genomics 67:40-47 (2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Palmer S.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Eye;  
 RX MEDLINE=22398257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Search completed: April 20, 2004, 01:36:59  
Job time : 22 secs

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavanti T.L., Scheetz T.E.,  
 Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
 Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettunen M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
 C -1- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex  
 (by similarity).  
 C -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.  
 C -1- TISSUE SPECIFICITY: Expressed in heart, kidney, liver, lung,  
 pancreas, and placenta, but not in skeletal muscle.  
 C -1- SIMILARITY: Contains 1 F-box domain.  
 C -1- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 -----  
 EMBL; AF176699; AAF03699.1; -  
 EMBL; AF174590; AAF04511.1; -  
 EMBL; AF199355; AAF09247.1; -  
 EMBL; AL022395; CAB37981.1; -  
 EMBL; BC055010; AAH55010.1; -  
 Genew; HGNC:13601; FBXL4.  
 MIM; 605654; -  
 InterPro; IPR001810; F-box.  
 InterPro; IPR007089; LRR cys.  
 InterPro; IPR006553; LRR cys sub.  
 InterPro; IPR008945; Skp1\_Skp2.  
 Pfam; PF00646; F-box; 1.  
 SMART; SM00256; FBOX; 1.  
 SMART; SM00367; LRR CC; 1.  
 PROSITE; PS00181; FBOX; 1.  
 Ubl conjugation pathway; Repeat; Leucine-rich repeat; Nuclear protein.  
 DOMAIN 277 332 F-BOX.  
 LRR 1.  
 REPEAT 400 421 LRR 1.  
 REPEAT 425 449 LRR 2.  
 REPEAT 450 475 LRR 3.  
 REPEAT 478 503 LRR 4.  
 REPEAT 504 524 LRR 5.  
 REPEAT 532 558 LRR 6.  
 REPEAT 559 583 LRR 7.  
 REPEAT 584 609 LRR 8.  
 CONFLICT 5 5 F -> Y (IN REF. 1).  
 CONFLICT 42 42 S -> G (IN REF. 3).  
 CONFLICT 122 122 Q -> K (IN REF. 1).  
 CONFLICT 126 126 Y -> I (IN REF. 1).  
 CONFLICT 219 219 E -> G (IN REF. 1).  
 CONFLICT 232 232 P -> A (IN REF. 1).  
 SEQUENCE 621 AA; 70096 MW; D8FD51A5C4F922D3 CRC64;  
 Query Match 4.0%; Score 126.5; DB 1; Length 621;  
 Best Local Similarity 21.4%; Pred. No. 0.047;  
 Matches 98; Conservative 65; Mismatches 161; Indels 135; Gaps 23;  
 121 QYFDCLSKLHFR-----RMIVKSDSLQNLARDGRGHVLHAKLKDCKSGFTDGLFHIGRECK 176  
 190 QFKCIQINPTNLIIEVNSLSLEYTETDLDAVLHGKVDKPVLSKTSILDMND---- 245

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 01:09:15 ; Search time 28 seconds  
(without alignments)

2081.859 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180

Sequence: 1 TKTSAPFLTLRLSNNTTE.....QRSDFPDVVPLDTATCVDT 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78:\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2126.5	66.9	592	T52139	LRR-containing F-b
2	818	25.7	585	E85040	hypothetical prote
3	812.5	25.6	594	T48087	transport inhibito
4	749.5	23.6	577	F86261	F13K23.7 protein -
5	693	21.8	614	T09902	hypothetical prote
6	229.5	7.2	618	T48193	hypothetical prote
7	209.5	6.6	712	C71419	hypothetical prote
8	188	5.9	607	E96598	protein F20N2.2 li
9	187.5	5.9	628	E84649	probable glucose r
10	186	5.8	489	S44609	hypothetical prote
11	176.5	5.6	467	G96837	unknown protein T2
12	176.5	5.6	690	T08604	hypothetical prote
13	172.5	5.4	518	D96512	hypothetical prote
14	169.5	5.3	563	T41312	probable DNA excis
15	168.5	5.3	522	T45861	hypothetical prote
16	159.5	5.0	656	E84547	hypothetical prote
17	154	4.8	463	T39987	probable regulator
18	152	4.8	381	T43444	hypothetical prote
19	149.5	4.7	678	B84856	hypothetical prote
20	148	4.7	554	B85072	hypothetical prote
21	147.5	4.6	276	T52349	F-box protein FB12
22	147.5	4.6	568	F86291	hypothetical prote
23	146	4.6	360	A96799	hypothetical prote
24	140	4.4	360	B63347	hypothetical prote
25	139	4.4	630	BWTT8Q	regulatory protein
26	139	4.4	630	S13724	ESAG 8 protein - T
27	135.5	4.3	934	T05201	hypothetical prote
28	134.5	4.2	1184	T71436	hypothetical prote
29	134.5	4.2	1301	D85188	disease resistance

30	133	4.2	630	2	A36359	VSG expression sit
31	132	4.2	565	2	A25226	RAD7 protein - yea
32	131.5	4.1	301	2	E85358	hypothetical prote
33	131.5	4.1	1217	2	T52348	disease resistance
34	127.5	4.0	1151	2	A41529	GRR1 protein - yea
35	126.5	4.0	1405	2	T04426	hypothetical prote
36	126	4.0	551	2	T23345	hypothetical prote
37	125	3.9	307	2	A85069	N7-like protein [i
38	124.5	3.9	250	2	T08680	hypothetical prote
39	122.5	3.9	1778	2	AF1116	internalin protein
40	122	3.8	1215	2	H84513	probable disease r
41	121.5	3.8	556	2	T33367	hypothetical prote
42	121.5	3.8	578	2	E96838	hypothetical prote
43	121	3.8	465	2	D96567	F6D8.13 [imported]
44	120.5	3.8	1189	2	T52346	disease resistance
45	119.5	3.8	419	2	H96695	hypothetical prote

## ALIGNMENTS

### RESULT 1

T52139  
LRR-containing F-box protein [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein At2g39940  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 16-Feb-2001  
C:Accession: T52139; C84823  
R:Xie, D.X.; Reys, B.F.; James, S.; Nieto-Rostro, M.; Turner, J.G.  
Science 280, 1091-1094, 1998  
A:Title: Coll: An Arabidopsis gene required for jasmonate-regulated defense and fertility  
A:Reference number: Z25980; MUID:98248619; PMID:9582125  
A:Accession: T52139  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-592 <X1E>  
A:Cross-references: EMBL:AF036340; PIDN:AAC17498.1  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-592 <STO>  
A:Cross-references: GB:AB002093; NID:g2088647; PIDN:AAB95279.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: Coll; At2g39940  
A:Map position: 2

Query Match 66.9%; Score 2126.5; DB 2; Length 592;

Best Local Similarity 69.6%; Pred. NO. 6.9e-159;

Matches 407; Conservative 72; Mismatches 97; Indels 9; Gaps 3;

QY 19 EERNVKTTRV-----VDVWLCVTPYDDPKDRDAVSQVCRWYELDSLTKGHVIALCY 73

Db 2 EDDPIKCKLSVATVDVDDVIBQVMTYITDPKDRDSASLVCRWFKIDSETRHVTMALCY 61

QY 74 TTTTARLRRRPHLESLLKKGKPRAMENLIPEDWGGHVTWVKESIQYFCDLCKSLHRR 133

Db 62 TATPDRLSRRPFLNLSKLGKPRAMENLIPENWGGVITPWVTEISNNLQKLSVHRR 121

QY 134 MIVKDSLDLQNLARDGHVHLAKLDKCSGFTTDDGLFHGTFRCKSLRVLFLEESSILEKDG 193

Db 122 MIVSDDLDRAKRADDETLKDKCSGFTTDLGLSVTHCRKIKTLMESSFSKDG 181

QY 194 EMLHELALNNVTINFLYTDIAVVKTEDELLAKNCPNLVSVKLTDCETLIDLVNFKH 253

Db 182 KWLHQLAQNHTSLEVLNFMTEFAKISPKDLETTARNCRSLVSVKVGDFEILELVGFFKA 241

QY 254 ASALBEFCGGTYNEE---PERYSALSPLAKLRLGLTYIGKNELPIVFMFAVLKLDLL 310

Db 242 AANLEFCGSLNEDIGMPKYMVLVPRKLCRLGLSYMGNEPILFFPAAQIRKLDLL 301  
QY 311 YAMLDTHDHCMLIQRCPNLEVRNIGRGLVGRCCCKRLKRLRIERGDDDDQMEDE 370  
Db 302 YALLEDETHCTLIQCPNLEVRNIGRGLVGRCCCKRLKRLRIERGADQMEDE 361  
QY 371 EGVSHRGLIALSGCSELEFMAVYVSDITNASLEHIGTHLKNICDFRLVLLDHEEKITD 430  
Db 362 EGLVSQGLIALAQCGCELEFMAVYVSDITNESLESIGTYLKNICDFRLVLLDHEEKITD 421  
QY 431 LPLDNGVRLALRGCDKLRRLFALYLRGGTLDVGLYIGQYSPNVRMMLLGVVGSDEGLL 490  
Db 422 LPLDNGVRLSLIGCKLRRLFALYLRGGTLDVGLYIGQYSPNVRMMLLGVVGSDEGLM 481  
QY 491 EFAGKCPSLQKLEMGCLFFSFERALAAVAATQTSRLYLWVGYSVPSGRDLLVWAPFW 550  
Db 482 EFSRGCPNLQKLEMGCC-FSERALAAVAATQTSRLYLWVGYSVPSGRDLLVWAPFW 540  
QY 551 NIELIPSKVATNNPDVTVVVHPAHILAYSLAGQSDPDDTV 595  
Db 541 NIELIPSRVPEVNOQGEIREMEHPAHILAYSLAGQSDPDDTV 585

RESULT 2  
E85040  
hypothetical protein AT4G03190 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: E85040  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; PMID:20083488; PMID:10617199  
A:Accession: E85040  
A:Status: preliminary  
A:Molecule type: DNA  
A:Cross-references: 1-585 <STO>  
A:Introns: 156/2; 320/3  
A:Gene: AT4G03190  
A:Map position: 4

Query Match 25.7%; Score 818; DB 2; Length 585;  
Best Local Similarity 34.6%; Pred. No. 1e-56;  
Matches 200; Conservative 109; Mismatches 237; Indels 32; Gaps 12;  
QY 32 VLDCVPIYDDPKORDAVSQVCRWYELDSLTKRHVTIALCYTTTPARLRPPHLESUK 91  
Db 9 VLEHILSFIDSNEGRNSVSVCKSWFETERKTRKRVFVGCYAVSPAATRRFPPEMRSLT 68  
QY 92 LKGPRAAMFNLIPEDWGHHVTPWKEISQYFDCLSKLHFRMIVKDSLOQLA---RDR 148  
Db 69 LKGPRAADFNLPDGGGYYVFWIEAIAAKSSSEIRMKRMVTVDECKEIAAASFKD- 127  
QY 149 GHVLHAKLDCSGFTDGLFHIGRCKSLRVLFEESSILEKDGWELHALLNNVTLET 208  
Db 128 ---FKVLVLTSCRGFTDGLAAIAATCRLNRLVLELRECEIVEDLGGDWLSYFPSSSTLSV 184  
QY 209 LNFVLTDAVVKIELELAKNCPNLVSVKLTDCETIID-LVNFHFKASALEEFCGTYNE 267  
Db 185 LDFSCLD-SEVKISDLERLSRSPNLKSLNPAVTLGLVSLRCAPQLTELGTGSFAA 243  
QY 268 E--PERYSATSLPAKLCR-----LGLTYIGKNELPIVFMFAAVLKLDLYAMLDTEHCH 320  
Db 244 QLKPEAPSKLSEAFNSCKQLQSLGLMDVLPYLPALYVCPGLTSLNLSYATVRPDLV 303  
QY 321 MLQRCPNLEVRNIGRGLVGRCCCKRLKRLRIERGDDDDQMEDEEGTVSHRGLI 380  
Db 304 ELURRCKLQKLMWMDIEDKGLAVASYCKELRELAVFSEPD--LDATNIPTEGLV 361  
QY 381 ALSGCSSELEFMAVYVSDITNASLEHIGTHLKNICDFRLVLLD--HEEKITDPLDNGVR 438

Db 362 FVSKGCRKLESVLYFCVQFTNAALFTIARKRPNLKCRLCVIEFPADPYKTNPDLKGF 421  
QY 439 ALARGCDKLRRLFALYLRGGTLDVGLYIGQYSPNVRMMLLGVVGSDEGLLFAKGCPS 498  
Db 422 AIAEGCDLRRLSV---SGLSDKAFYIGKHAKYRMLSLAFAGSDSLMLHLLSGCES 478  
QY 499 LQKLEMGCLFFSFERALAAVAATQTSRLYLWVGYSVPSGRDLLVWAPFWNIELIPSR 558  
Db 479 LKLEIRDCP-FGDTALLEHAALKETMRSLWMSSCFVSGACKLLSQMPRLANVEVI--- 534  
QY 559 KVATNNPDVTVVVHPAHILAYSLAGQSDPDDTV 595  
Db 535 ---DEHPPSPRESSPVERIYIYRTVAGPRMDTPEFV 568

RESULT 3  
T48087  
transport inhibitor response protein TIR1 [imported] - Arabidopsis thaliana  
N:Alternate names: protein T20010.80  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 03-Nov-2000  
C:Accession: T48087; T51946  
R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Le  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24484  
A:Accession: T48087  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-594 <OBE>  
A:Cross-references: EMBL:AJ163816  
A:Experimental source: cultivar Columbia; BAC clone T20010  
R:Ruegger, M.; Dewey, E.; Gray, B.; Hobbie, L.; Turner, J.; Estelle, M.  
submitted to the EMBL Data Library, May 1997  
A:Description: The TIR1 protein of Arabidopsis functions in auxin response and is relat  
A:Reference number: Z25881  
A:Accession: T51946  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-594 <RUE>  
A:Cross-references: EMBL:AF005047; PIDN:AA69175.1  
C:Genetics:  
A:Gene: TIR1  
A:Map position: 3  
A:Introns: 156/2; 320/3  
A:Note: T20010.80

Query Match 25.6%; Score 812.5; DB 2; Length 594;  
Best Local Similarity 34.8%; Pred. No. 1e-55;  
Matches 201; Conservative 104; Mismatches 250; Indels 23; Gaps 11;  
QY 32 VLDCVPIYDDPKORDAVSQVCRWYELDSLTKRHVTIALCYTTTPARLRPPHLESUK 91  
Db 13 VLEHVSFIQDKDRNSVSVCKSWFETERMCRKRVFIGNCYAVSPATVIRFPKRVSYE 72  
QY 92 LKGPRAAMFNLIPEDWGHHVTPWKEISQYFDCLSKLHFRMIVKDSLOQLAARDGHV 151  
Db 73 LKGPRAADFNLPDGGGYYVFWIEAIAAKSSSEIRMKRMVTVDECKEIAAASFKN- 131  
QY 152 LHALKLDKCSGFTDGLFHIGRCKSLRVLFEESSILEKDGWELHALLNNVTLETNF 211  
Db 132 FKVLVLTSCRGFTDGLAAIAATCRLNRLVLELRESDVDDVSGHLSHPDDTVTSVSLNI 191  
QY 212 YLTDIAVVKIELELAKNCPNLVSVKLTDCETIID-LVNFHFKASALEEFCGTYNE-- 268  
Db 192 SCL-ASEVSFSALERLVTRCPNLKSLNRAVPELKLATLQAPQLEELGTGGYTAERV 250  
QY 269 PERYSATSLPAKLCR-----LGLTYIGKNELPIVFMFAAVLKLDLYAMLDTEHCH 323  
Db 251 PDVYSGLSVALSCCKELRCLSGFWDVAVPAYLVAVSVCSRLTTLNLSYATVQSDVAVKLL 310  
QY 324 QRCPNLEVRNIGRGLVGRCCCKRLKRLRIERGDDDDQMEDEEGTVSHRGLIALS 383  
Db 311 CQCPKLRWLVDYIEDAGLEVLAACKDLRELVR--FFSEPFVMEPNVALTEQGLVSVS 368

QY 384 QGCESEYMAVYVSDITNASLEHIGTHLKNLCDFRLVLLDHE--EKITDLPDNGVRALL 441  
 DB 369 MGCKPESLVYFCQMTNAAITITARNRNMTFRCLIEPKADYLTLEBDDIGFGAIV 428  
 QY 442 RGCCKLRRLRYLRGGITDVLGYIGQYSPNVWRMLLGVVGSDDAGLLFFAKCPSLQK 501  
 DB 429 EHCKDLRLSL--SGLLTQVFEYIGTYAKMELMSVAFAGSDLDLGMHVLGSCDSLK 485  
 QY 502 LEMRGCLFFSERALAVAATQLTSLRYLWQVGYVSPSGRDLVLMARPPNLELIPSKVA 561  
 DB 486 LEINDCP-FGDKALLANASKLETRSLWMSSCSVSGACKLLGQKPKLVNVEIDERG-A 543  
 QY 562 TINTNPDETUVVVEHPAHILAYYSLAGORSDFPDVTVPLD 599  
 DB 544 PDSRPESCPV---ERVFIYRTVAGPRDFMPGFWNMMD 577  
 RESULT 4  
 F13K23.7 protein - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Nov-2001  
 C:Accession: F86261  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, X.; Liu, X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: F86261  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-577 <STO>  
 A:Cross-references: GB:AE005172; NID:g8698729; PIDN:AAF78487.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 Query Match 23.6%; Score 749.5; DB 2; Length 577;  
 Best Local Similarity 33.8%; Pred. No. 8.8e-51;  
 Matches 197; Conservative 97; Mismatches 258; Indels 31; Gaps 13;  
 QY 30 DVVLCVPIYDDPKORDAVSOVCRWRVYELDSLTKHVTIALCYTTTPARLRERPHLES 89  
 DB 6 DEVIEHVFDFVASHKDRNSISLVCKSWHKIERFSRKEVFIGNCAINPERLIRFPCLKS 65  
 QY 90 LKLGKPRAAFNLPEDWGGHVTTPWKEISQYFCLSKSLHPRMIVKDSIQNLARDRG 149  
 DB 66 LTLKGPFPADNLVPHENGGFVHPNIEALASRVGLLELRKRVVTVDESLLLSRFA 125  
 QY 150 HVLHAKLDKCSGFTDGLFHFGRCKSLRVLPFLESSILEKDGWHLHSLANNVTLET 209  
 DB 126 N-FKSLVIVSCGFTDGLASIAANCRHLRLDLQNEIDHRGQWLNCFPDSCITLMSL 184  
 QY 210 NYFL-----TDIAVVKIEDLEAKNCPNLVSKLTDCEITLD-LVNPFFKHASALEBFCGT 264  
 DB 185 NFACLKGETNVA-----LERLVARSPLNLSKIANRVLPLDALRUMSCAPQLVDLGVS 239  
 QY 265 YNEEPERYSAISPAKLCRL-----GLTYIGKNELPITVFMFAAVLKDLILYAMLDPTE 317  
 DB 240 YENEPPEPSAKLNTAKKITSLSRUSGFLEVAPLCFAPYPCQNLSINLSYAAEIOG 299  
 QY 318 DHCM-LIQRCPNLEVLTRNVIGDRGLVGRCKLRIRIERGDDDDQGMEDDEGTSH 376  
 DB 300 NHLIKLIQCKRLQRLWILSDIGKGLAVVAATCKELQLRV-FPSDVHGEEDNNASVTE 358  
 QY 377 RGLIALSQGCESELEYMAVYVSDITNASLEHIGTHLKNLCDFRLVLLD-HE-EKITDLPD 434

DB 359 VGLVAISAGCPKLHSILYFCQMTNAAILAVAKNCPNIFRCLIEPHEKPHDHTIFQSLD 418  
 QY 435 NGVRALLRGCDKLRRLRYLRGGITDVLGYIGQYSPNVWRMLLGVVGSDDAGLLFFAK 494  
 DB 419 EGFGAIVQVACKGLRLSLV--SGLLTQVFEYIGTYAKMELMSVAFAGSDLDLGMHVLN 475  
 QY 495 GCPSLQKLEMRGCLFFSERALAVAATQLTSLRYLWQVGYVSPSGRDLVLMARPPNLEL 554  
 DB 476 GCKQKRLIRIRERGDDDDQGMEDDEGTSHRGLIALSQGCESELEYMAVYVSDITNASL 404  
 QY 555 IPSRK--VAINTNPDETUVVVEHPAHILAYYSLAGORSDFPDVTV 595  
 DB 535 INENNINGMEQNEDEKVD---KLYLYRTVTVGTRKDAPPV 574  
 RESULT 5  
 T09902  
 hypothetical protein T22A6.220 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999  
 C:Accession: T09902  
 R:Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mc  
 submitted to the Protein Sequence Database, June 1999  
 A:Reference number: Z16896  
 A:Accession: T09902  
 A:Molecule type: DNA  
 A:Residues: 1-614 <BEV>  
 A:Cross-references: EMBL:AL078637; GSPDB:GN00062; ATSP:T22A6.220  
 A:Experimental source: cultivar Columbia; BAC clone T22A6  
 C:Genetics:  
 A:Gene: ATSP:T22A6.220  
 A:Map position: 4  
 A:Introns: 192/2; 358/3  
 Query Match 21.8%; Score 693; DB 2; Length 614;  
 Best Local Similarity 30.7%; Pred. No. 2.6e-46;  
 Matches 190; Conservative 108; Mismatches 277; Indels 44; Gaps 13;  
 QY 1 TKTSAFFLFTLSLSNMTHERNVRT-----RVVDVVLDCVPIYDDPKORDAVSOVCR 54  
 DB 19 SSSSSAATFT-----NKSRSFKSPPPCPDHVLENVLENTLOFLTSCRDNVSLVCR 71  
 QY 55 RWYELDSLTRKHVTIALCYTTTPARLRERPHLESKLKGPRAAFNLPEDWGGHVTTP 114  
 DB 72 SWYRVEATRLVEVIGNCISLSPARKLIRFRKRVSLVKGPRPADFNLMPPNWAQFSP 131  
 QY 115 WKXISQYFDCILSLHPRMIVKDSIQNLARDRGHVLHAKLDKCSGFTDGLFHFGRF 174  
 DB 132 VVAATAKAYPWLEKVLKRMFVTDLLALLA-ESPPGFKELTLVCEGFGTSGIAIVANK 190  
 QY 175 CKSLRVLPFLESSILEKDGWHLHSLANNVTLETNLFYLTDAVVKIEDLELAKNCPNL 234  
 DB 191 CRQLKVLDMESVTDDELWDWISCFPEGETHLESLSFDCVE--SPINFALEELVRSPL 249  
 QY 235 VSVKLTLD-CEITLDLVNFFKHASALEBFCGTYN-----BEPERYSAISPAK--LCRL 284  
 DB 250 KKLIRNFRVSEELHRLMVRAPOLTSLGTSFSPDNVPQGEQPDYAAAFACKSIVCLS 309  
 QY 285 GLTYIGKNELPITVFMFAAVLKDLILYAMLDPTEHMLIORCPNLEVLTRNVIGDRGLE 344  
 DB 310 GFREFRPVLLAISVCANLSTNFSYANISPHMLKPIISNCHNIRVFWALDSIRDEGLQ 369  
 QY 345 VLGRCKLRIRIERGDDDDQGMEDDEGTSHRGLIALSQGCESELEYMAVYVSDITNASL 404  
 DB 370 AVAATCKELRLIRIFFPDP---REDSEGFVSGVGLQISEGCRKLESILYFQNTINGAV 426  
 QY 405 EHGTHLKNLCDFRLVLL-----DHEEKTLDLPNGVRALLRGCDKLRRLRYLRGGI 459  
 DB 427 TAMSNCPLQTVFRILCIRMGRRHPD---VTGKPMDDGFGAIVNCKKLTFLAV---SGLL 480  
 QY 460 TDVGLGYIGQYSPNVWRMLLGVVGSDDAGLLFFAKCPSLQKLEMRGCLFFSERALAVAA 519  
 DB 481 TDEAFSYIGEYKLIIRTLISVAFAGNSDKALRVYLEGCPKLQKLEIRDSP-FGDVGLRSGM 539





[illegible]

140 VLS-----LAVGCKKLRRLHRLCINVASVEWLEVFGLKLEELC-----180  
275 ISLPKLCRLGLTYIGKNEL-----PIVFMFAVLKLLDLYAMLDTEHCLMI 323  
181 -----IKNCRA-----IGEGDLIKLRNSWRKLTSLQFEVDANYRMYK-VYDQLOVERWPKQL 231  
324 QRCNPVLETRNVI--GDRGLEVLGRCKRLKRLRIERGDDDDQGDMEDEBEGTVSHRGLIA 381  
232 VPCDSIVELSLGNCITAPGRGLACVLNCKNLEKHL---DMCTGVSDSD-----IIA 281  
382 LSQCCSELEVMAYVSDITNASLEHIGTHLKNLCDFRLVLLDHEKHTDLPDNGVRALL 441  
282 LVQASHLSISLRVPS-----DFTLPLANN-----ITLRLTDESLSAIA 321  
442 RGCDKLRFPALYLRG-----GLTDVGLGYIQYSPNFMWMLLGVYGE-SDAGLLEFAK 494  
322 QHCKSLKSPKISFSDGEGFSLFTLQGIITLQKCP-VRELSLDHVCVFNDMG-MEALC 379  
495 GCPSLQKLEMRGCLFSEERALAVAATOLTSRLVLMW-QYGVSPSPGRDILLVMAPPFNIE 553  
380 SAQKLEILELVHCOEVSDEGL-ILVSQFSPSLNVLKSLKCLGVTDDG-----MRP-----427  
554 LIPSRKVATNTNPDETVWVE 573  
428 LVGSHKL-----ELLVVE 440

RESULT 12  
T08604  
hypochemical protein GRR1 - soybean  
C:Species: Glycine max (soybean)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999  
C:Accession: T08604  
R:Chen, W.; Atherly, A.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z15438  
A:Accession: T08604  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-690 <CHE>  
A:Cross-references: EMBL:AF019910; NID:g2407789; PID:g2407790  
A:Experimental source: variety L85-3044; root  
C:Genetics:  
A:Gene: grr1

Query Match 5.6%; Score 176.5; DB 2; Length 690;  
Best Local Similarity 20.6%; Pred. No. 1.1e-05;  
Matches 123; Conservative 76; Mismatches 170; Indels 227; Gaps 26;  
QY 7 FLFTLSRS-----NMTEERNVRKTRVVDVVLDC--VIFYIDDPKORDAVSQVCR 55  
DB 81 FLFRGLASVFLHLSAENGLSRSRQSRTSIESLPDECLFEILRLPAGQDRSVCASVSKR 140  
QY 56 WYELDSLTRKHVTIATCYTTTPARLRRRFPHLSSKLKXKPRAMFNLIPEDNHGHVTPW 115  
DB 141 WLMLLS-----SICKT-----EHSYSGSTGNEN-----163  
QY 116 VKEISQVFDCLSLHFRMIVKDSLDQLNLA--RDRGHVLHAKLDKCS--GFTTDLGF 169  
DB 164 -QEISDRGYLSRLEGGK--ATDVRLAIAVGRASRG-LGKLTIRGNSDRGVTVNGLK 219  
QY 170 HIGRFCKSLRVLFEESSILEKCEWELHMLANNVTLETNLYTDTAVKIEDELLAK 229  
DB 220 AIAHGCSFLKCSLWDVATVGDVG--LIEIASGCHQLEKLD--LCKCFNISKDTLIAVAK 275  
QY 230 NCPNLVSKLTDCEILDVNFVFKHASALEEFCGTYNEEPERYSAISLPKLCRLGLTYI 289  
DB 276 NCPNLAEELS-----284  
QY 290 GKNELPIVFMFAVLKLLDLYAMLDTEHCLMIQRCNLEVLNTRNIGRGLVLRGRC 349  
DB 285 -----IENCPN-----IGNEGLQAIGK- 301

137 KQSDQLNARDRGHVLHAKLDKCSGFTTDLGFHIGRFCKSLRVLFLEESSILEKDGWL 196  
163 TDASCENLGR-YCHKLYNLNLENCSSITDRAMKYIGDGCNLSYLNISWCDALQDRGV--219  
197 HELALNN-TVLETFNY----LTDIAVVKTEDELLAKNCPNLVSVKLTDCETILDVNF 251  
220 -QIILSNCKSLDTLILRGCEGLTENVFGSVEAHMGAIKLNLQCLQTLDTITQNIAN--276  
252 KHASALEEFCGTYNEEPERYSAISLPKLCRLGLTYIGRKNELPIVFMFAVLKLLGLLY 311  
277 -GATALEYLCNMCNQISDR-----SLVSLG-----301  
312 AMLDTEHCLMIQRCNLEVLTR--NVIGDRGLEVLGRCKRLKRLRIERGDDDDQGMED 369  
302 -----QHSHNLKVLSESGCTLLGDNQFIPLARGCQRLRDME-----339  
370 BEGTVSHRGLIALSQGSELEYMAV-VVSDITNASLEHIGT-HLKNLDCDFRLVLLDHEK 427  
340 DCSLISDHTINSANNCTALRELSHCELITDESIGNLASKHRETL--NVLELDNCPQ 396  
428 ITDPLDNGVRALLRGCDKLRFPALY-----LRGGITDVGGLGYIQYSP-----472  
397 LTDSTLSH-----LEHCKALKRIDLYDQNVSKEAIVRFQVVDVAVLSYDCIYFPCSTRS 451  
473 NVRWMLLGVVGESD 486  
452 TENWYKLTATKSD 465

RESULT 11  
G96837  
unknown protein T21f11.10 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: G96837  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: G96837  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-467 <STO>  
A:Cross-references: GB:AB005173; NID:g6730728; PIDN:AAF27118.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T21f11.10  
A:Map position: 1

Query Match 5.6%; Score 176.5; DB 2; Length 467;  
Best Local Similarity 23.6%; Pred. No. 6.4e-06;  
Matches 132; Conservative 69; Mismatches 186; Indels 171; Gaps 29;  
QY 45 DRDAVSQVCRWYELDSLTRKHVTIATCYTTTPA-----RLRRPPLHLSKLGKPRAA 99  
DB 21 DRNLSLSCKRFFSLDNQORSIRIG-C-GLVPASDALLSLCRFPNLSKVEI-----71  
QY 100 MENLIPDWGCHVTWPVKIEISQYFCLKSLHFRMIVKDSLDQLNARDRGHVLHAKLDK 159  
DB 72 -----IYSGWMSKLGQVD-----DQGLLVLTN-CHSLTDLTSLF 106  
QY 160 CSGFTTDLGFHIGRFCKSLRVLFLEESSILEKDGWELHMLANNVTLETNLF--YLTIDIA 217  
DB 107 CTFITDVGIGHLSS-CPLESSL-----KLNFAPIITGG 139  
QY 218 VKKIEDELLAKNCPNLVSVKLTDC--BILDVNFVFKHASALEEFCGTYNEEPERYSA 274





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 01:04:00 ; Search time 72 Seconds  
(without alignments)  
2655.612 Million cell updates/sec

Title: US-10-009-791-22

Perfect score: 3180

Sequence: 1 TKTSAFLFTLSLRNMTEE.....QRSDPDTVVPLDTATCVDT 606

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2126.5	66.9	592	10 O04197	O04197 arabidopsis
2	1855.5	58.3	630	10 Q8W0G4	Q8W0G4 oryza sativ
3	1766.5	55.6	589	10 Q84QA7	Q84QA7 oryza sativ
4	1200	37.7	237	10 Q8H6H7	Q8H6H7 glycine max
5	819	25.8	585	10 Q9C5Y7	Q9C5Y7 arabidopsis
6	818	25.7	585	10 Q9ZRL2	Q9ZRL2 arabidopsis
7	813	25.6	585	10 Q94AU0	Q94AU0 arabidopsis
8	812.5	25.6	594	10 Q24660	Q24660 arabidopsis
9	760.5	23.9	575	10 Q9LW29	Q9LW29 arabidopsis
10	749.5	23.6	577	10 Q9LPW7	Q9LPW7 arabidopsis
11	741.5	23.3	619	10 Q9LTX2	Q9LTX2 arabidopsis
12	722.5	22.7	575	10 Q7XVM8	Q7XVM8 oryza sativ
13	720	22.6	635	10 Q9AUH6	Q9AUH6 populus tre
14	693	21.8	614	10 Q9STV5	Q9STV5 arabidopsis
15	693	21.8	623	10 Q8RWQ8	Q8RWQ8 arabidopsis
16	636.5	20.0	587	10 Q9FRB6	Q9FRB6 oryza sativ

17	621	19.5	603	10 Q8H7P5	Q8H7P5 oryza sativ
18	256.5	8.1	610	10 Q8LGK0	Q8LGK0 arabidopsis
19	249.5	7.8	610	10 Q9C5D2	Q9C5D2 arabidopsis
20	235	7.4	395	10 Q9SRR1	Q9SRR1 arabidopsis
21	229.5	7.2	618	10 Q9M004	Q9M004 arabidopsis
22	229	7.2	665	10 Q8RWU5	Q8RWU5 arabidopsis
23	221	6.9	642	10 Q8GZ31	Q8GZ31 arabidopsis
24	217	6.8	628	10 Q94GB3	Q94GB3 oryza sativ
25	217	6.8	628	10 Q7XGT5	Q7XGT5 oryza sativ
26	213	6.7	405	10 Q9FMW7	Q9FMW7 arabidopsis
27	213	6.7	772	5 Q9VF10	Q9VF10 drosophila
28	209.5	6.6	712	10 Q23399	Q23399 arabidopsis
29	203	6.4	154	10 Q82085	Q82085 lycopersico
30	193.5	6.1	2159	5 Q86IU5	Q86IU5 dictyosteli
31	188	5.9	601	10 Q9ARM5	Q9ARM5 arabidopsis
32	188	5.9	607	10 Q9ZWC6	Q9ZWC6 arabidopsis
33	187.5	5.9	628	10 Q9SKK0	Q9SKK0 arabidopsis
34	184	5.8	522	10 Q7X785	Q7X785 oryza sativ
35	183.5	5.8	466	5 Q8T3G0	Q8T3G0 caenorhabdi
36	180	5.7	479	10 Q7XV84	Q7XV84 oryza sativ
37	180	5.7	482	10 Q7XVD0	Q7XVD0 oryza sativ
38	179	5.6	436	4 Q961G2	Q961G2 homo sapien
39	176.5	5.6	467	10 Q9M8M4	Q9M8M4 arabidopsis
40	176.5	5.6	690	10 Q22512	Q22512 glycine max
41	175	5.5	422	11 Q9CZV8	Q9CZV8 mus musculu
42	175	5.5	423	4 Q9NVQ8	Q9NVQ8 homo sapien
43	175	5.5	423	11 Q8BHL6	Q8BHL6 mus musculu
44	172.5	5.4	440	10 Q7XVD2	Q7XVD2 oryza sativ
45	172.5	5.4	464	5 Q9V605	Q9V605 drosophila

## ALIGNMENTS

RESULT 1

ID	O04197	PRELIMINARY;	PRT;	592 AA.
AC	O04197;			
DT	01-JUL-1997 (TrEMBLrel. 04, Created)			
DT	01-JUL-1997 (TrEMBLrel. 04, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Hypothetical protein (AT2G39940/T28M21.10).			
GN	T29M21.10 OR COI1			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	[1]_TaxID=3702;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Rounsley S.D., Ketchum K.A., Lin X., Phillips C.A., Brandon R.C.,			
RA	Fuhrmann J.L., White O., Kerlavage A.R., Adams M.D., Somerville C.R.,			
RA	Venter J.C.;			
RL	Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	PubMed=12244256;			
RA	Feys B.J., Benedetti C.S., Penfold C.N., Turner J.G.;			
RL	Plant Cell 6:751-759(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=98248619; PubMed=9582125;			
RA	Xie D.X., Feys B.P., James S., Nieto-Rostro M., Turner J.G.;			
RT	"COI1: an Arabidopsis gene required for jasmonate-regulated defense			
RT	and fertility.";			
RL	Science 280:1091-1094(1998).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,			
RA	Bowser L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,			

RA Hayaehizaki Y., Ishida J., Jiang P.X., Jones T., Jones T., Kamiya A.,  
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu X.X.,  
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
RA Tang C.C., Tortumi M., Yamada K., Yu G., Yu S., Shinozaki K.,  
RA Davis R.W., Theologis A., Ecker J.R.,  
RT "Arabidopsis cDNA clones.",  
RN Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RL  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,  
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayaehizaki Y.,  
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
RA Tang C.C., Tortumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,  
RT "Arabidopsis ORF clones.",  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF002109; AB95279.1; -  
DR EMBL; AF036340; AAC17498.1; -  
DR EMBL; AY045625; AAK73383.1; -  
DR EMBL; AY133556; AAM91386.1; -  
DR PIR; T52139; T52139.  
SQ SEQUENCE 592 AA; 67665 MW; 1DDCF04990144C06 CRC64;  
  
Query Match 66.9%; Score 2126.5; DB 10; Length 592;  
Best Local Similarity 69.4%; Pred. No. 6.6e-173;  
Matches 407; Conservative 72; Mismatches 97; Indels 9; Gaps 3;  
  
QY 19 EERNVKTIV-----VDVVLDDCVIPYIDDKDQDAVSQVCRWYELDSLTKHVTALCY 73  
Db 2 EDDPIKCKLSCVATVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDDVDD 61  
  
QY 74 TTPPAIRRRFPFLESGLKGGKRAAMFNLIPEDDWGHVTPWVKEISQVDCIKSLHFR 133  
Db 62 TATPDLRFRFPNLSILKGGKRAAMFNLIPEDDWGHVTPWVKEISQVDCIKSLHFR 121  
  
QY 134 MIVKSDQLNARDGRHVLHALKDKSGFTTDLGFIHGRFCKSLRVLFLEESSILEKOG 193  
Db 122 MIVSDLDLRLAKARADDLTLKDKSGFTTDLGFIHGRFCKSLRVLFLEESSILEKOG 181  
  
QY 194 EWLHEALANNTVLETNFYLTADVVKIELELLAKNCNPLSVKLTDCHEILDVNFEEKH 253  
Db 182 KWLHEALANNTVLETNFYLTADVVKIELELLAKNCNPLSVKLTDCHEILDVNFEEKH 241  
  
QY 254 ASALEEFCCGTYNEE---PERYSAISLPALCRGLTYIGKNELPIVFMFAAVLKLDL 310  
Db 242 AANLEEFCCGSLNEDIGMPEKYNLVFPRLCRGLSYMGPNEMPILFPAAQIRKLDL 301  
  
QY 311 YAMLDTEHCKLTCORCNPLVETRNVDGRGLEVLGRCCRLKRLRIERGDDDDQGWED 370  
Db 302 YALLETEHCKLTCORCNPLVETRNVDGRGLEVLGRCCRLKRLRIERGDDDDQGWED 361  
  
QY 371 EGVTVSHRGLTALSGCSELYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITD 430  
Db 362 EGVTVSHRGLTALSGCSELYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITD 421  
  
QY 431 LPLDNGVRLLRGCDKLRLRFPALYLRGGLTDVGLGYIGQYSPNVRWMLLGVGESDAGLL 490  
Db 422 LPLDNGVRLLRGCDKLRLRFPALYLRGGLTDVGLGYIGQYSPNVRWMLLGVGESDAGLL 481  
  
QY 491 EFAKGCPSLOKLEMRGCLFSEALAVAATQTSRLYLMVQGVSPSGRDLVWAPFW 550  
Db 482 EFSRGCPNQLKLEMRGCC-FSEALAAAATVTLPSRLYLMVQGVSPSGRDLVWAPFW 540  
  
QY 551 NIELIPSRKVATNTPDVTWVHPAHILAYYSLAGORSDFPDV 595  
Db 541 NIELIPSRKVATNTPDVTWVHPAHILAYYSLAGORSDFPDV 585

RESULT 2  
Q8W0G4

ID Q8W0G4 PRELIMINARY; PRT; 630 AA.  
AC Q8W0G4;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE P0529E05.15 protein.  
GN P0529E05.15.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]\_TaxID=4530;  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RC Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
RT clone:P0529E05.",  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF003279; BAB84399.1; -  
DR Gramene; Q8W0G4; -  
DR InterPro; IPR001810; F-box.  
DR InterPro; IPR001611; LRR.  
DR Pfam; PF00646; F-box; 1.  
DR Pfam; PF00560; LRR; 1.  
SQ SEQUENCE 630 AA; 70619 MW; BFA12F8F1C93F23C CRC64;  
  
Query Match 58.3%; Score 1855.5; DB 10; Length 630;  
Best Local Similarity 60.9%; Pred. No. 1.1e-149;  
Matches 352; Conservative 87; Mismatches 132; Indels 7; Gaps 5;  
  
QY 28 VDVVLDDCVIPYIDDKDQDAVSQVCRWYELDSLTKHVTALCYTTPAIRRRFPFHL 87  
Db 56 VPDEALHLVGHVEDPRDEAASVCRWRHRIADALTRKHVTAFVCAAPARLRERFPL 115  
  
QY 88 ESLLKGGKRAAMFNLIPEDDWGHVTPWVKEISQVDCIKSLHFRMIKSDQLNARD 147  
Db 116 ESLSLGGKRAAMFNLIPEDDWGHVTPWVKEISQVDCIKSLHFRMIKSDQLNARD 175  
  
QY 148 RGHVLHALKDKSGFTTDLGFIHGRFCKSLRVLFLEESSILEKOGELHEALANNTVLE 207  
Db 176 RGHVLHALKDKSGFTTDLGFIHGRFCKSLRVLFLEESSILEKOGELHEALANNTVLE 235  
  
QY 208 TLFNYLTADVVKIELELLAKNCNPLSVKLTDCHEILDVNFEEKHSALEEFCCGTYNE 267  
Db 236 TLFNYLTADVVKIELELLAKNCNPLSVKLTDCHEILDVNFEEKHSALEEFCCGTYNE 294  
  
QY 268 --EPERYSAISLPALCRGLTYIGKNELPIVFMFAAVLKLDLTYAMLDTEHCKMLIQ 325  
Db 295 VGLTYKVKVFPRLCRGLTYMGPNEMPILFPAAQIRKLDLTYAMLDTEHCKMLIQ 354  
  
QY 326 CNLNVLETRNVIGDRGLEVLGRCCRLKRLRIERGDDDDQGWEDEGTVSHRGLIALSQ 385  
Db 355 CNLNVLETRNVIGDRGLEVLGRCCRLKRLRIERGDDDDQGWEDEGTVSHRGLIALSQ 414  
  
QY 386 CSELEYMAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITDPLDNGVRLLRGCD 445  
Db 415 CRELEYIAAVYSDITNASLHIGTHLKNLCPRLVLLDHEEKITDPLDNGVRLLRGCD 474  
  
QY 446 KLRRFPALYLRGGLTDVGLGYIGQYSPNVRWMLLGVGESDAGLLFAKGCPSLOKLEMR 505  
Db 475 KLRRFPALYLRGGLTDVGLGYIGQYSPNVRWMLLGVGESDAGLLFAKGCPSLOKLEMR 534  
  
QY 506 GCLFFSEALAVAATQTSRLYLMVQGVSPSGRDLVWAPFWNIELIPSRKVATN-- 563  
Db 535 SCC-FSEALSLAVLQMPISRLYLMVQGVSPSGRDLVWAPFWNIELIPSRKVATN-- 593  
  
QY 564 TNPDETVVHPAHILAYYSLAGORSDFPDV 601  
Db 594 TEDGEPVDH-AQVLAAYYSLAGORSDFPDV 630  
  
RESULT 3



01-MAR-2003 (TREMELrel. 23, Last sequence update)  
01-MAR-2003 (TREMELrel. 23, Last annotation update)  
Putative coronatine-insensitive 1 (Fragment).  
COI1  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid1 I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. TGX1835-2E;  
RA Christiansen L.C., Ulvskov P.;  
RT "Putative Glycine max sbCOI1 homologous to Arabidopsis COI (atFBL2)." atFBL2.1;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF536527; AN31713.1; -.  
DN  
FT NON TER 1  
FT NON TER 237  
SQ SEQUENCE 237 AA; 26280 MW; 6F8C6917AF74E1FD CRC64;

Query Match 37.7%; Score 1200; DB 10; Length 237;  
Best Local Similarity 99.1%; Pred.No. 2.9e-94;  
Matches 230; Conservative 1; Mismatches 1; Indels 0; Gaps

Qy 374 VSHRGIALISQCSELEYMAVYVS DITNASLEHIGTHLKNCDFRLVLLDHEEKITDLPL  
Db 6 VSHRGIALISQCSELEYMAVYVS DITNASLEHIGTHLKNCDFRLVLLDHEEKITDLPL  
Qy 434 DNGVRALLGCDKLRFFALYIRGGITDVGLGYICQSPNVRWMLLGYYGSDAGILLEFA  
Db 66 DNGVRALLGCDKLRFFALYIRGGITDVGLGYICQSPNVRWMLLGYYGSDAGILLEFA  
Qy 494 KGCPSLQKLEMGRCLFFSERALAVAATOLTSLRYLVWGQYGVSGRDLVVWAPPFWNIE  
Db 126 KGCPSLQKLEMGRCLFFSERALAVAATOLTSLRYLVWGQYGVSGRDLVVWAPPFWNIE  
Qy 554 LIPSRKVATNPDETVVVEHPAHILAYYSLAGQRSDPDTVPPLDTATCYD 605  
Db 186 LIPSRKVATNPDETVVVEHPAHILAYYSLAGQRSDPDTVPPLDTATMCD 237

RESULT 5  
ID Q9CSY7 PRELIMINARY; PRT; 585 AA.  
AC Q9CSY7;  
DT 01-JUN-2001 (TREMELrel. 17, Created)  
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DB GSRI-like protein 1.  
GN GRH1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid1 II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Thelander M., Fredriksson D., Schouten J., Hoge H.C., Ronne H.;  
RT "Cloning by pathway activation in yeast: identification of an  
Arabidopsis thaliana F-box protein that turns on glucose repression.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF291816; AK011147.1; -.  
DN InterPro; IPR001810; F-box.  
DR InterPro; IPR001611; LRR.  
DR Pfam; PF00646; F-box; 1.  
DR Pfam; PF00560; LRR; 1.  
DR SMART; SM00256; FBOX; 1.  
SQ SEQUENCE 585 AA; 65678 MW; CAF4E7C2B4150BDD CRC64;

Query Match 25.8%; Score 819; DB 10; Length 585;  
Best Local Similarity 34.6%; Pred.No. 4e-61;  
Matches 200; Conservative 109; Mismatches 237; Indels 32; Gaps

QY 32 VLDVPIVDDPKDQDAVQVCRWYELDSLTRKHVTIALCYTTTPARLRERFFPLESLK 91  
 Db 9 VLEHILSPIDNEDRNSVSLCKSWFETERKTRKRVFVGNVCYAVPAVTRFFPEMRSLT 68  
 QY 92 LKGPRAAMENLIPEDGHHVTPWVKELSOYFDCIKSLHFRMIVKDSLDQNLA---RDR 148  
 Db 69 LKGPHFADYNLVPDGGGYAWPWEIAWAKSSLEIRMKRMVVTDECKEIAASFKD- 127  
 QY 149 GHVLHAKLDCSGFTTDLGPHIGRCKSLRVLFLSESSILEKDGEMHLEALANTVLET 208  
 Db 128 ---FKVLVTSCGFTDGIATAATCRNLRLVLEFCEIVEDIGGDMLSYFPESTSLVS 184  
 QY 209 LNFYLTDAVVKIEDLELAKNCPLNSVKLTDCETLD-INVFFKHASALEEFCGGTYNE 267  
 Db 185 LDFSCLD-SEVKISDLERLVSRSPLKSLKPNVTLGLVSLRCAPQTELGTGSPFAA 243  
 QY 268 E--PERYSALSPAKLCR---LGLTYIGKNELPIVFMFAAVLKKLDLILYAMLDTHC 320  
 Db 244 QLKPEAFSKLSEAFNSCKQLSLGLWDLVPEYLPALYSVCPGLTSLNLSYATVRMPDLV 303  
 QY 321 MLQRCNLEVLTRNVIGDRGLVGRCKRLKRLRIERGDDQDQMEDEEGTVSHRGLI 380  
 Db 304 ELLRRCCKLQKLVWMDLIEDKLEAVASYCKELRLRVFPSEPD--LDATNIPLEQGLV 361  
 QY 381 ALSQGCSELEVMAYVSDITNASLEHIGTHLKNLDCDFRLVLLD--HEEKITDPLDNGVR 438  
 Db 362 FVSKGCKRLSVLFCVQFTNAALFTIARKRNLCFRLCVIEPAPDYKTNFPLDKGFK 421  
 QY 439 ALLRGCDLRRPALVLRGGLTVGLGIVGYQVSNVWMLLGYVGSDEAGLEFAKGCPS 498  
 Db 479 LKLEIRDCP-FGDTALLEHAALKETRMSLWMSFCVSPGACKLISQKMPRLNVEVI--- 534  
 QY 559 KVATNTPDETIVVVEHPA-HILAYYSLAGORSDFPDV 595  
 Db 535 ---DEHPESPRESSPVERIYRTVAGPRMDTPEFV 568

## RESULT 6

Q9ZR12 ID Q9ZR12 PRELIMINARY; PRT; 585 AA.  
 AC Q9ZR12  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Putative homolog of transport inhibitor response 1 (Putative F-box protein AtFBL16)  
 DE protein AtFBL16  
 GN F4C21.11 OR AT4G03190.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 [1]  
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA Habermann K., de la Bastide M., Huang E.N., Gnoj L., Schutz K.,  
 RA Preston R., Calma C., Martensen R., Parnell L.D., Dedhia N.,  
 RA McCombie W.R.  
 RT "Arabidopsis thaliana BAC F4C21 from chromosome IV near 17 cM."  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,  
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,  
 RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,  
 RA Mewes H.W., Lemcke K., Meyer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RN SEQUENCE FROM N.A.  
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,  
 RA Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,  
 RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,  
 RA Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,  
 RA Ecker J.R., Theologis A.;  
 RT "Arabidopsis Open Reading Frame (ORF) Clones";  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC005275; AAD14447.1; -;  
 DR EMBL; AL161496; CAB77804.1; -;  
 DR EMBL; AY150427; AAN12969.1; -;  
 DR PIR; E85040; E85040.  
 DR InterPro; IPR001810; F-box.  
 DR InterPro; IPR001611; LRR.  
 DR Pfam; PF00646; F-box; 1.  
 DR Pfam; PF00560; LRR; 1.  
 DR SMART; SM00256; FBOX; 1.  
 SQ SEQUENCE 585 AA; 65647 MW; D39D627C82864D83 CRC64;

Query Match 25.7%; Score 818; DB 10; Length 585;  
 Best Local Similarity 34.6%; Pred. No. 4.8e-61;  
 Matches 200; Conservative 109; Mismatches 237; Indels 32; Gaps 12;

QY 32 VLDVPIVDDPKDQDAVQVCRWYELDSLTRKHVTIALCYTTTPARLRERFFPLESLK 91  
 Db 9 VLEHILSPIDNEDRNSVSLCKSWFETERKTRKRVFVGNVCYAVPAVTRFFPEMRSLT 68  
 QY 92 LKGPRAAMENLIPEDGHHVTPWVKELSOYFDCIKSLHFRMIVKDSLDQNLA---RDR 148  
 Db 69 LKGPHFADYNLVPDGGGYAWPWEIAWAKSSLEIRMKRMVVTDECKEIAASFKD- 127  
 QY 149 GHVLHAKLDCSGFTTDLGPHIGRCKSLRVLFLSESSILEKDGEMHLEALANTVLET 208  
 Db 128 ---FKVLVTSCGFTDGIATAATCRNLRLVLEFCEIVEDIGGDMLSYFPESTSLVS 184  
 QY 209 LNFYLTDAVVKIEDLELAKNCPLNSVKLTDCETLD-INVFFKHASALEEFCGGTYNE 267  
 Db 185 LDFSCLD-SEVKISDLERLVSRSPLKSLKPNVTLGLVSLRCAPQTELGTGSPFAA 243  
 QY 268 E--PERYSALSPAKLCR---LGLTYIGKNELPIVFMFAAVLKKLDLILYAMLDTHC 320  
 Db 244 QLKPEAFSKLSEAFNSCKQLSLGLWDLVPEYLPALYSVCPGLTSLNLSYATVRMPDLV 303  
 QY 321 MLQRCNLEVLTRNVIGDRGLVGRCKRLKRLRIERGDDQDQMEDEEGTVSHRGLI 380  
 Db 304 ELLRRCCKLQKLVWMDLIEDKLEAVASYCKELRLRVFPSEPD--LDATNIPLEQGLV 361  
 QY 381 ALSQGCSELEVMAYVSDITNASLEHIGTHLKNLDCDFRLVLLD--HEEKITDPLDNGVR 438  
 Db 362 FVSKGCKRLSVLFCVQFTNAALFTIARKRNLCFRLCVIEPAPDYKTNFPLDKGFK 421  
 QY 439 ALLRGCDLRRPALVLRGGLTVGLGIVGYQVSNVWMLLGYVGSDEAGLEFAKGCPS 498  
 Db 479 LKLEIRDCP-FGDTALLEHAALKETRMSLWMSFCVSPGACKLISQKMPRLNVEVI--- 534  
 QY 559 KVATNTPDETIVVVEHPA-HILAYYSLAGORSDFPDV 595  
 Db 535 ---DEHPESPRESSPVERIYRTVAGPRMDTPEFV 568

## RESULT 7

Q94AU0 ID Q94AU0 PRELIMINARY; PRT; 585 AA.  
 AC Q94AU0  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)





```

RN RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012187; AAF78487.1; -
DR EMBL; AF099541; AAM20393.1; -
DR EMBL; BT002118; AAN72129.1; -
DR PIR; F86261; F86261.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR008945; Skp1_Skp2.
DR Pfam; PF00646; F-box; 1.
SQ SEQUENCE 577 AA; 64906 MW; 8E419B4D80368661 CRC64;

Query Match      23.6%; Score 749.5; DB 10; Length 577;
Best Local Similarity 33.8%; Pred. No. 3.4e-55;
Matches 197; Conservative 97; Mismatches 258; Indels 31; Gaps 13;

QY 30 DVVLCVPIYDDPKDRDAVSQVCRRWYELDSLTKHVTIALCYTTTPARLRRFPFHL 89
Db 6 DEVIEHVDFVASHKDRNSISLVCKSWHKIERFSRKEVFIGNVCYAINPERLIRFPCLKS 65
QY 90 LKLGKPRAMNLIPEDWGGHVTWVKIEISQYFDCLSLHFRMTVVKDSLDQNLARDG 149
Db 66 LTLGKPFADPNLVPHWGGVHVPWIEALARSVGLBELRLKRWVVTDESLLDSRFA 125
QY 150 HVLHAKLDKCSGFTDGLFHIGRCKSLRVLFLEESSILEKDGWHLHNLNNTVLETL 209
Db 126 N-FKSLVLSVCSGFTDGLASIAANCRHLRELDLQNEIDHGGWLNCFDSCCTTMSL 184
QY 210 NFYL-----TDIAVKIEDELLAKNCPNLVSVKLTDCELD-LVNFFKHASALEEFCOGT 264
Db 185 NFACLKGETNVA-----LERLVARSPLKSLKLNRAVPLDALRLMSCAPQLVDLGVGS 239
QY 265 YNEEPERYSAISLPKLCRL-----GLTYGKNEPLIVFNFVAVLKLLDLYLMLDTE 317
Db 240 YNEPDPSFAKMTAKYTSLSRSLSGFLEVAFLCPAFYFICQNLISLNSLYAAEIQG 299
QY 318 DHCM-LIQCNPLEVLETRNVIGDRGLGVIGRCCRLKRLRIERGGDDQGMEDSGTVSH 376
Db 300 NHLIKLIQCKLQRLWILDSIGDKGLAVVAATCKELQELRV-FPSDVHGEEDNNAVTE 358
QY 377 RGLIALSGCSELEWYVSDITWASIEHIGTHLKNLCLDFRLVLD-HE-EKITDLPID 434
Db 359 VGLVAISGCPKLHSLYFCOMTNAALJAVAXNCPNFIREFLCILEPHKPDHITFQSLD 418
QY 435 NGVALLRGCKLRRFALYLRGGTLTDVGLGVIGYQVSPNRMILGYVYGESDAGLEPAK 494
Db 419 EGFGLVAVACKGLRLSV---SGLTDQVFLYIGMAEQLEMLSTAFAGDQKGLVYN 475
QY 495 GCPSLQKLEMRGCLFFSERALAVAAQTLSRLYLVQGVSPSGRDLVLMARPPFNIEL 554
Db 476 GCKMKRLEIRDSP-FGNAALLADGVRYETMRSLWMSCEVTIGGCKLAQNSPLNVEI 534
QY 555 IPSRK--VATNPNPETVVVHPAHILAYLSLAGORSFDPDTV 595
Db 535 INENENNGMEQNEEREKVD---KLTYLRYTVVGRKDPAPPV 574

RESULT 11
Q9LTX2 PRELIMINARY; PRT; 619 AA.
AC Q9LTX2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transport inhibitor response 1 protein (At5g49980/k9p8_12).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN RP SEQUENCE FROM N.A.
RA MEDLINE=20181125; PubMed=10718197;
RX STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63 (2000).
RN RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cdna clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Kong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cdna clones.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB024C32; BAA97019.1; -
DR EMBL; AY056431; AAL08287.1; -
DR EMBL; AY139774; AAM98092.1; -
DR EMBL; BT004536; AAO42782.1; -
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
SQ SEQUENCE 619 AA; 69316 MW; D0614AF071EE4FD2 CRC64;

Query Match      23.3%; Score 741.5; DB 10; Length 619;
Best Local Similarity 33.4%; Pred. No. 1.8e-54;
Matches 196; Conservative 100; Mismatches 255; Indels 35; Gaps 12;

QY 28 VDVVLCVPIYDDPKDRDAVSQVCRRWYELDSLTKHVTIALCYTTTPARLRRFPFHL 87
Db 54 VLENVLENVLQFLDSRCDRNAASLVCKSWHVEALTRSEVFIGNVCYALSPLRITQRFKRV 113
QY 88 ESKLKGKPRAMNLIPEDWGGHVTWVKIEISQYFDCLSLHFRMTVVKDSLDQNLARD 147
Db 114 RSLVLKGRPADFNLMPPDNGANFAPWVSTMAQAYPCLEKVDLKRMEVTTDDLLALA-D 172
QY 148 RGHVLHAKLDKCSGFTDGLFHIGRCKSLRVLFLEESSILEKDGWHLHNLNNTVLE 207
Db 173 SFGFKELILVCCSGFGTSGISIVANKRKLKVLIDIESEVTDDDEVDMISCFPEVDTCLE 232
QY 208 TLFNFLTIDIAVKIEDELLAKNCPNLVSVKLTDC-CEILDVNFFKHASALEEFCOGTY- 265

```

Qy	27	RVVVVLDCVITPYIDDDPKORDAVSQVCRRWYELDSLTRKHVTITACYTTTPARLRRFPFH	86
Db	68	QVLENVLENVLWFTTSRKDRNAASLWRSWRYRYEALTRSDLFTIGNCYAVSPKRAMSRFR	127
Qy	87	LESLSLKGKPRAAFMNLLIPEDWGSHVTWPWKEYSQVFDCLKSLHPRMTVKDSIDQLNLAR	146
Db	128	IRSVTLKGKPRAFDNLMPYPWGHAFAPWVSAMNTYPWLEKVHLKMSVTTDDDLALLAE	187
Qy	147	DRGHVHLHALKDKCGSETTIDGLPHIGRFCKSLRVLFLESSLKDGWELHALLNNTVL	206
Db	188	SPSG-FKELVLCVCEGFGTSGLAIVWSRCQLAKVLDLISDSVDDVDWISCFPDTETCL	246
Qy	207	ETLAFYLTDAVWKIEDLELLAKNCNLSVKLTD-CEILDLVNFPFKHASALEECGGTY	265
Db	247	ESLTFDCVD-CFIDFDELERLVARSPSLKKGLRNRYVSGQLYRLMIRAPHLTHLGTGSF	305
Qy	266	N-----EEPBRYSAISLPACL-CRLGLTYIGKNELPIVFMFAVLKGLDILLAYAMD	315









US 100097910FP1



Creation date: 06-03-2004  
Indexing Officer: SUK - SOPHY UK  
Team: OIPEBackFileIndexing  
Dossier: 10009791

Legal Date: 05-17-2004

No.	Doccode	Number of pages
1	SRNT	2

Total number of pages: 2

Remarks:

Order of re-scan issued on .....